TRANSCRIPTIONAL INITIATION OF SORGHUM MITOCHONDRIAL DNA

By

BIN YAN

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1995

UNIVERSITY OF FLORIDA LIBRARIES

ACKNOWLEDGEMENTS

I would like to express my deep appreciation to Dr. Daryl R. Pring, my major advisor, for his patient and expert guidance throughout the work. His constant encouragement and constructive criticisms during the course of this research and the preparation of this dissertation are gratefully acknowledged.

I am indebted to many members of our laboratory I met during my stay between August of 1990 and August of 1995, particularly to Drs. Hoang V. Tang and Figuhr Muza, for many helpful suggestions and discussions. I am thankful to members of my graduate committee, Drs. Prem Chourey, Christine Chase, Dean Gabriel and Corby Kistler, for their invaluable advice and instructions. Sincere gratitude is also extended to the faculty, staff members, and other people in the Fifield Hall for their assistance and friendship.

Special thanks go to my family for their understanding and support during my schooling years.

TABLE OF CONTENTS

PAGE

ACKNOWLEDGEMENTS ii

ABSTRACT v

LITERATURE REVIEW 1

Introduction 1
Mitochondrial Genomes 4
Animal Mitochondrial DNA 4
Fungal Mitochondrial DNA 5
Plant Mitochondrial DNA 6
Transcription Initiation in Mitochondria 13
Transcription Initiation of Animal Mitochondrial DNA 13
Transcription Initiation of Yeast Mitochondrial DNA 15
Transcription Initiation of Plant Mitochondrial DNA 16
Regulation of Mitochondrial Gene Expression 20

MATERIALS AND METHODS 23

Genetic Stocks 23
Isolation of Mitochondrial DNA, RNA and Protein 23
Recombinant DNA methods 27
Characterization of Initiated Transcript Termini 35
RNA Polymerase Assay 37

RESULTS 42

Transcription Initiation of *atp6-1* 43
Transcription Initiation of *atp6-2* 48
Transcription Initiation of *atp9* 54
Transcription Initiation of *orf107* 66
Transcription Initiation of *urf209* 71
Transcription Initiation of *orf265* 74
In vitro Transcription of *atp6-2* with Maize Extract 83

DISCUSSION 86
The Transcription Initiation Regions of *atp6-1* and *atp6-2* 88
The Transcription Initiation Region of *atp9* 89
The Transcription Initiation Regions of *orf107* 91
The Transcription Initiation Region of *urf209* 92
The Transcription Initiation Regions of *orf265* 93
In vitro Transcription of *atp6-2* 94
Consensus Motif of Sorghum mtDNA Promoters 95

REFERENCES 101
BIOGRAPHICAL SKETCH 113

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

TRANSCRIPTIONAL INITIATION OF SORGHUM MITOCHONDRIAL DNA

By

Bin Yan

August, 1995

Chairman: Dr. Daryl R. Pring
Major Department: Plant Pathology

A study of transcriptional initiation of sorghum mitochondrial DNA was conducted, including an examination of four sorghum mitochondrial genes (*atp6-1*, *atp6-2*, *atp9* and *urf209*) and two open reading frames associated with cytoplasmic male-sterility in sorghum (*orf265* and *orf107*). Transcriptional initiation sites were determined by primer extension analyses and *in vitro* capping with guanylyltransferase in combination with ribonuclease protection. Nine transcriptional initiation regions were identified, including two regions for *orf265* and three regions for *orf107*. In addition to a distal promoter, *orf265* has a second initiation region that resides in a maize-like sequence, and shares a promoter homologous to that of *T-urf13* in maize. The transcripts of *atp9* have two origins. The gene is cotranscribed with *trn*fM, which is 323 bp 5' to *atp9*. Processing of the precursor transcript at the 3' terminus of *trn*fM results in 5' processed transcript termini of *atp9*. Unusually, transcripts with termini 3' to *trn*fM are also suitable substrates for

guanylyltransferase, suggesting that a transcriptional initiation region of *atp9* exists 3' to *trn*fM. This novel observation suggests that as few as 21 bp 5' to the transcription initiation site of *atp9* may be sufficient for recognition by RNA polymerase. Alternatively, part of the *trn*fM coding region may be involved as a promoter of *atp9* and play a role in the initiation of transcription.

Sequences surrounding the nine transcription initiation sites were compared. Five of the nine initiation regions, including initiation regions of obligate genes *atp6-2* and *atp9*, contain the proposed YRTA element found in other plant mitochondrial systems. Repeated elements G(A/T)3-4 or C(A/T)3-4, present just upstream of the transcription initiation sites, as found in several plant systems, were observed for some sorghum genes in this study.

The transcription initiation region of *atp6-2* shares high similarity with a maize *atp6* promoter, and was tested in a heterologous *in vitro* transcription system with partially purified maize mitochondrial RNA polymerase. Transcription of *atp6-2* initiated at a predicted site.

LITERATURE REVIEW

Introduction

Mitochondria, as a site of energy metabolism, contribute uniquely to a cell among cellular organelles. In addition to the principal function of oxidative phosphorylation, mitochondria are also involved in the biosynthesis of numerous key intermediates in cellular metabolism, such as pyrimidines, amino acids, nucleotides, phospholipids, folate coenzymes, urea, heme, and contribute to import and localize proteins, phospholipids and a few tRNAs to the proper positions (Attardi and Schatz 1988).

Like chloroplasts, mitochondria have their own genomes and possess a separate enzymatic machinery to express their genetic information into functional proteins (Tzagoloff and Myers 1986). The mitochondrial energy-generating pathway is composed of five enzyme complexes (NADH dehydrogenase, succinate dehydrogenase, cytochrome c reductase, cytochrome c oxidase and ATP synthase). Apart from the conventional phosphorylating pathways, mitochondria of all plant species examined have a nonphosphorylating alternative pathway (Day et al. 1980, Siedow and Berthold 1986). In general, mitochondrial genomes may contain genes coding for several subunits of Complex I (*nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, and *nad7*), one gene coding for one subunit of Complex III (*cob*), three genes coding for the three largest subunits of complex IV (*cox1, cox2*, and *cox3*), several genes coding for several subunits of the ATP synthase

complex (*atp1*, *atp6* and *atp9*), and some genes coding for tRNAs as well as several subunits of ribosomal RNAs (Bonen and Brown 1993, Tzagoloff and Myers 1986).

Mutations in the mitochondrial genome can be limited lesions that affect single enzymatic functions which result in mitochondrial defects, or can be lethal to a cell (Attardi and Schatz 1988). In human mitochondria, mutations of mtDNA resulting in either amino acid substitution, nucleotide substitution in tRNA genes; insertion-deletion or copy number changes of mitochondrial DNA (mtDNA) have been found to be associated with some common degenerative diseases, including neurogenic muscle weakness, ocular myopathy, and maternally inherited myopathy and cardiomyopathy (Wallace 1992). In yeast, mutations in mitochondrial DNA can cause a respiratory deficiency, resulting in the inability to utilize nonfermentable substrates (Costanzo and Fox 1990). Plant mitochondrial DNA mutations have been demonstrated to be associated with cytoplasmic male sterility (CMS) (Pring and Lonsdale 1989, Braun et al. 1992), abnormal growth mutants (Newton et al. 1990), the teosinte cytoplasmic-associated miniature mutation (Cooper et al. 1990) and a leaf variegation phenotype in *Arabidopsis* (Martinez-Zapater et al. 1992). Consequently, proper function of mitochondria, i. e. proper gene expression, is essential.

One of the major aspects of study on mitochondrial gene expression and regulation is the characterization of DNA elements involved in the transcription initiation of mtDNA, which in all cases, except one (Chang and Clayton 1986b) are located around transcription initiation sites (Biswas et al. 1990, Clayton 1991, Rapp and Stern 1992). Two general approaches have been widely used to identify transcript termini of mitochondrial DNA. First, 5' ends of *in vivo* RNAs can be identified by either primer extension, S1 nuclease mapping, direct RNA sequencing or *in vitro* capping with guanylyltransferase, or a combination of these techniques. Guanylyltransferase catalyzes the incorporation of GMP from [α-^{32}P]GTP into 5' termini which carry di- or tri-phosphates and results in a cap

structure, whereas processed termini, which contain a 5' monophosphate, will not be a suitable template for the enzyme (Christianson et al 1983). Since mitochondrial RNAs are not capped *in vivo*, they can be capped *in vitro* by this enzyme.

Second, the putative functional promoter sequences can be defined by testing the cloned mtDNA as templates in an *in vitro* transcription system with purified homologous RNA polymerase. These two approaches, especially *in vitro* transcription analysis, have been exploited to great detail in studying the mechanisms of transcription initiation of animal and yeast mtDNAs.

Compared with their counterparts in animal and yeast, plant mitochondrial genomes are much larger and more complex (Gray 1989, Lonsdale and Grienenberger 1992). Frequent recombination and rearrangement events could result in duplications of part or entire regions of protein-coding genes or generations of chimeric configurations, including transcriptional regulatory elements. Therefore, it is expected that multiple transcription initiation sites would exist in plant mtDNA (Gray et al. 1992).

Transcript-mapping studies have indicated that there can be multiple sites of transcription initiation in plant mtDNA (Binder and Brennicke 1993b, Brown et al. 1991, Covello et al. 1991, Hanic-Joyce and Gray 1991, Mulligan et al. 1988a, 1988b, Rapp and Stern 1992). Utilizing *in vitro* transcription systems, wheat mitochondrial *cox2* (Hanic-Joyce and Gray 1991) and maize *atp1*, *atp6* and *cox3* promoters were defined (Rapp and Stern 1992). The functional analysis of the maize *atp1* promoter demonstrated that a conserved 11 nucleotide motif of (A/T)CRTA(G/T)A(A/T)AAA (R; purine) is an essential element for the *atp1* promoter function (Rapp and Stern 1992, Rapp et al. 1993). Nevertheless, despite the general good agreement of many mapped 5' transcript termini with the proposed YRTA (corresponding to the initially proposed CRTA, Y; pyrimidine) motif, the initiation regions for some transcripts show little similarity to the conserved YRTA

element (Binder et al. 1994, Lizama et al. 1994, Mulligan et al. 1991). The uncertainty regarding the nature of plant mitochondrial promoter emphasizes the need for characterization of more promoters of plant mtDNAs.

Mitochondrial Genomes

It is widely accepted that the mitochondrion originated from alpha-purple bacterium as an endosymbiont (Attardi and Schatz 1988). As far as the structure, gene organization and the mode of expression are concerned, mitochondrial DNA exhibits an extraordinary diversity in different organisms of the eukaryotic kingdoms of animals, plants and fungi (Gray 1989). This diversity probably reflects a polyphyletic origin of past endocytotic events or a monophyletic origin that has led to diversity of present-day mitochondria through different evolutionary pathways in different eukaryotic cells (Attardi and Schatz 1988).

Animal Mitochondrial DNA

The most intriguing feature of the vertebrate mitochondrial genome is its extremely compact organization. With few exceptions, vertebrate mtDNAs are very similar in size (16.5-17.6 kb), and present as circular molecules at high copy number at a level of 1000-10000 mtDNA molecules per somatic cell (Clayton 1991). The genomes have an invariant gene order, and are organized and expressed in an essentially identical manner (Anderson et al. 1981). Except for a short DNA segment surrounding the origin of replication, vertebrate mtDNA is saturated by structural genes, which are adjacent one to another (Clayton 1984, 1991). In most cases, there is a tRNA gene between adjacent protein coding genes, and all of the genes lack introns (Clayton 1984, 1991). Most of the genes are encoded by the heavy (H)-strand, including 14 tRNA genes, 2 rRNA genes and 13

protein coding genes that encode seven subunits of NADH dehydrogenase, apocytochrome b of ubiquinol cytochrome c reductase, three subunits of cytochrome c oxidase and two subunits of ATP synthase (*atp6* and *atp8*) (Gray, 1989).

While the little size variation among vertebrate mtDNAs is mainly confined to the noncoding D-loop region, which contains the origin of DNA replication as well as promoters for two-strand transcription (Clayton 1991), the size variation over threefold among invertebrate resulted from absence or duplication of the coding and noncoding regions (Gray 1989). In the nematode worm *Ascaris suum*, the mitochondrial genome lacks the *atp8* gene, and the genome size is as small as 14.3 kb (Wolstenholme et al. 1987). The sea scallop *Plactopecten magellanicus* mtDNA has large direct tandem duplications of the genome, and the genome size extends to 32.9 kb (Zouros 1992). Mitochondrial gene order varies among vertebrates, insects, echinoderms, nematode and platyhelminths (Gray 1989). Nevertheless, animal mitochondrial genomes, both vertebrate and invertebrate mtDNAs, maintain an extremely compact arrangement.

Fungal Mitochondrial DNA

The gene structure and organization of fungal mtDNA, which varies from 17.6 kb to 115 kb among ascomycetes (Gray 1986), contrast remarkably with vertebrate mtDNAs. From observations with electron microscopy, pulsed field gel electrophoresis and fluorescent microscopic technique, linear DNA molecules were proposed as the major form of mitochondrial genomes in fungi (Bendich 1993). The large size variation of fungal mtDNAs is due not only to the intergenic spacer length but also to the presence or absence of optional introns.

In contrast to animal mtDNAs in which no introns have been found, several fungal mitochondrial genes, such as *cob*, *cox1* and *rrn21* in *Saccharomyces cerevisiae*, are

themselves discontinuous. Some of the introns contain open reading frames coding for protein (maturase) involved in the intron splicing (Costanzo and Fox 1990). Other features, including duplication of portions of the genome, presence of short repetitive sequences and the number of short repetitive sequences, and presence of novel open reading frames, also contribute to the size variations of fungal mitochondrial genomes (Gray 1989, Tzagoloff and Myers 1986).

Generally, fungal mtDNA codes for several major protein components of respiratory chain enzymes, including subunit I, II and III of cytochrome c oxidase, apocytochrome b, subunit 6, 8 and 9 (which is nuclear-encoded in animal mtDNAs, Tzagoloff and Myers 1986) of ATP synthase, and a full set of tRNA as well as the two rRNAs (Tzagoloff and Myers 1986). In addition, fungal mtDNAs encode several low abundant proteins which may play an important role in the transposition of the introns, such as reverse transcriptase in yeast (Kennell et al. 1993).

Plant Mitochondrial DNA

Organization and genes

Higher plant mitochondrial genomes, which vary in size from 200 kb to approximately 2500 kb, are much larger and more complex than their animal or fungal counterparts (Lonsdale 1989, Lonsdale and Grienenberger 1992). Even the smallest plant mitochondrial genome, which is 208 kb in *Brassica hirta*, is larger than the largest fungal mitochondrial genome, which is 176 kb in *Agaricus bitorquis*, and larger than most chloroplast genomes (Gray 1989). While several unique features contribute to the genome expansion, the most interesting character is the relative high frequency of rearrangements that involve both intra- and intermolecular recombinations (Lonsdale and Grienenberger

1992, Pring and Lonsdale 1985). Other features include presence of repeated sequences (coding or noncoding) ranging from a few up to several hundred base pairs, and possession of integrated chloroplast sequences (Lonsdale and Grienenberger 1992, Pring and Lonsdale 1985).

The process of recombination between repeated DNA elements, large repeated sequences as well as small nucleotide repeats (Andre et al. 1992), results in either sequence inversions or the resolution of circular molecules into smaller subgenomic species (Lonsdale and Grienenberger 1992). This process could occur within molecules, as well as between different circular DNA molecules, which share sequence homology, and thus give rise multimeric forms of individual circle and more complex patterns of the genome (Lonsdale 1989, Lonsdale and Grienenberger 1992). A "master circle" model was proposed for plant mitochondrial genome, in which the entire genetic complexity is contained in a circular chromosome that may be resolved into subgenomic molecules by recombination through directly repeated sequences (Lonsdale 1989, Lonsdale and Grienenberger 1992, Schuster and Brennicke 1994). Based on observations of pulsed field gel electrophoresis and electrophoresis of moving pictures of DNA released by lysis of organelles, linear DNA molecules were proposed as the major *in vivo* form of mitochondrial genomes in higher plants (Bendich 1993).

Apart from a high-molecular-weight DNA component that represents the main genomic DNA, in many plant species, mitochondria also contain some smaller circular and linear DNAs (Lonsdale and Grienenberger 1992, Pring and Lonsdale 1985). While most of these plasmid sequences contain no sequences required for mitochondrial functions, the 2.3 kb linear plasmid of maize contains the plasmid genes for *trn*P and *trn*W (Leon et al. 1989, Marechal et al. 1987). These plasmid DNAs are sometimes integrated into the

genomic mtDNAs, and contribute to the complexity and large size of plant mitochondrial genomes (Lonsdale and Grienenberger 1992).

Promiscuous chloroplast DNA sequences are found in mitochondrial DNA (Stern and Lonsdale 1982), contributing to the large size of the genome. While most of the plastid sequences appear to be nonfunctional pieces of genes, several tRNAs seem to be transcribed from sequences of plastid origin in the mitochondrial compartment (Lonsdale 1989).

Most of the standard set of mitochondria genes found in mammalian and fungal mtDNAs have also been localized in many plant mitochondrial genomes. In addition, plant mitochondrial genomes have several genes that are not detected in fungal or animal mtDNAs. These genes include *rrn5*, *atp1*, *nad7* (Bonen et al. 1994), *nad9* (Lamattina et al. 1993), an ABC-type (ATP-binding cassette) heme transporter gene (Bonnard and Grienerbenger 1995, Jekabsons and Schuster 1995), *rps12* (Gualberto et al. 1988) and *rps14* in some plant mtDNAs (Newton 1988).

Unlike mammals (Anderson et al. 1981) and fungi (Heckman and RajBhandary 1979, Martin et al. 1977), plant mitochondrial genomes do not code for sufficient tRNAs to support protein synthesis (Dietrich et al. 1992). Analysis of tRNAs and the tRNA genes revealed the tRNAs whose genes are missing in plant mitochondrial genomes are actually imported from the cytosol (Marechal-Drouard et al. 1988).

A number of plant mitochondrial native tRNA genes have been sequenced, and it was found that the plant mitochondrial tRNAs resemble eubacterial and chloroplast tRNAs (65-80% sequence similarity) rather than mitochondrial tRNAs from fungi or animals at the level of both primary sequence and potential secondary structure (Dietrich et al. 1992). The maturation of the plant mitochondrial tRNA (Marchfelder and Brennicke 1994) is similar to that of yeast (Costanzo and Fox 1990), where 5'-leader and 3'-trailer sequences are

removed by the endonuclease RNAse P (Altman 1981) and a 3'-endonuclease respectively, followed by addition of CCA at the 3' end .

In several plant mitochondrial genomes, *trn*fM or pseudo *trn*fM is found 5' close to protein coding genes (Krishnasamy et al. 1993, Sangare et al. 1989, Singh and Brown 1991), such as *atp6* (Singh and Brown 1991). The gene coding for subunit 6 of ATP synthase is located 3' to *trn*fM in normal radish and fertile *Brassica napus*, and 3' to a truncated pseudo *trn*fM in Polima cytoplasm of *B. napus*. In Ogura cytoplasm of radish, an open reading frame, *orf138*, was found 3' to *trn*fM. In maize, one of the two copies of *trn*fM was revealed 5' to a transcribed protein gene (Sangare et al. 1989).

Transcript mapping with primer extension analysis of the downstream protein coding genes demonstrated that the 5' end of the shortest transcript of those genes were mapped precisely to the 3' end of *trn*fM (Singh and Brown 1991). A role of a tRNA or tRNA-like element in the formation of 5' transcript terminus of the downstream-located gene was proposed, whereby the tRNA is cotranscribed with the protein coding gene and a precise endonucleolytic cleavage of the precursor results in a mature 3' end of the tRNA, and the 5' transcript termini of the protein coding gene (Singh and Brown 1991).

Introns have been described for several plant mitochondrial genes. Among angiosperm mtDNAs, 22 different introns (Thomson et al. 1994) have been detected, including *cox2*, and 18 introns of genes *nad1*, *nad2*, *nad4*, *nad5* and *nad7* (Bonen et al. 1994, Wahleithner et al. 1990). These introns, ranging between 700 and 3400 nucleotides, are another factor that enlarge the plant mitochondrial genome.

Unlike the yeast mitochondrial genome, in which no editing event has been reported yet, RNA editing (Covello et al. 1989, Gualberto et al. 1989, Hiesel et al. 1989) is, probably, ubiquitous in the mitochondria of flowering plants, both monocotyledonous and dicotyledonous species (Bonnard et al. 1992, Gray et al. 1992, Schuster and Brennicke

1994). Editing occurs predominantly within coding regions, and most are C-to-U substitutions (Bonnard et al. 1992, Gray et al. 1992, Schuster and Brennicke 1994). Transcripts of virtually every protein-coding gene found in plant mitochondria, except *T-urf13* (Ward and Levings 1991) and *atp6* in some lines of radish (Krishnasamy and Makaroff 1994), are edited (Pring et al. 1993b) The net effect of C-to-U RNA editing in plant mitochondria is to make encoded proteins more similar in sequence to their nonplant homologs (Bonnard et al. 1992, Gray et al. 1992, Schuster and Brennicke 1994).

Cytoplasmic male sterility

In addition to the expected complement of genes whose products are involved in electron transport and ATP synthesis, plant mitochondria also contain a large number of unidentified open reading frames and chimeric genes, including genes associated with cytoplasmic male sterility in plants (Braun et al. 1992).

Cytoplasmic male sterility (CMS) is maternally inherited, results in pollen abortion, but does not affect female fertility (Levings 1993). This trait is of significant value agronomically in hybrid seed production by eliminating manual emasculation. With few exceptions (Chen et al. 1993), mitochondrial DNA mutations have been demonstrated to be associated with CMS trait, and the mitochondrial genome rearrangements associated with CMS in several plant species have been identified and characterized (Braun et al. 1992, Hanson et al. 1991, Pring and Lonsdale 1989).

The best studied system is in maize Texas cytoplasm (cms-T). The cms-T cytoplasm was extensively used to produce maize hybrids in the past. In 1969 and 1970, fungal diseases, Southern corn leaf blight and yellow corn leaf blight, caused by *Bipolaris maydis* race T and *Phyllosticta maydis*, respectively, occurred which severely blighted maize carrying cms-T (Braun et al. 1992, Hanson et al. 1991, Pring and Lonsdale 1989).

Thus, cms-T carried CMS factors as well as factors for disease susceptibility. The toxins produced from those two pathogens were especially virulent to the cms-T maize. An unusual gene, *T-urf13*, was associated with the CMS and toxin sensitivity traits (Dewey et al. 1986).

T-urf13 arose by rearrangements involving both intramolecular and intermolecular recombinational events (Dewey et al. 1986). The *T-urf13* coding region contained significant similarity to the 3' flanking and coding sequences of the *rrn26* gene, and the 5' flanking region of *T-urf13* showed high similarity to the 5' flanking region of *atp6* (Dewey et al. 1986). In addition to the recombinant fragment, the cms-T mitochondrial genome carried complete, functional copies of *rrn26* and *atp6* located elsewhere. An open reading frame, designated *orf25*, was downstream immediately from *T-urf13* (Dewey et al. 1986).

T-urf13 was expressed in cms-T male-sterile maize as a 13 kD polypeptide (Dewey et al. 1987, Wise et al. 1987b). While in cms-T maize restored to fertility, URF13 levels were significantly reduced (Dewey et al. 1987, Forde et al. 1978, Forde and Leaver 1980) by the nuclear gene *Rf1* alone (Dewey et al. 1987), although both *Rf1* and *Rf2* were required to restore fertility. *Rf1* affected expression of *T-urf13* by processing some of the *T-urf13* transcripts post-transcriptionally (Dewey et al. 1987, Kennell et al. 1987, Kennell and Pring 1989).

Revertants derived from tissue culture were male-fertile and toxin-insensitive (Brettell et al. 1980, Gengenbach et al. 1977, Wise et al. 1987a). Expression of *T-urf13* in heterologous systems of *E. coli* (Dewey et al. 1988), yeast (Glab et al. 1990, Huang et al. 1990), insect (Korth and Levings 1993) and tobacco (Chaumont et al. 1995, von Allmen et al. 1991) demonstrated that *T-urf13* was responsible for the toxin sensitivity and strongly associated with cms trait. A model has been proposed that URF13 forms inner mitochondrial membrane channels in the presence of toxin or anther-specific compound(s)

(Braun et al. 1992, Levings 1993), resulting in the permeability of the inner mitochondrial membrane, and thus led to disease sensitivity, and probably, pollen abortion (Braun et al. 1992).

In petunia, one type of cms has been described that can be restored by a single dominant *Rf* allele. The cms determinant was localized to a region of mtDNA that contained an open reading frame, *pcf* (Boeshore et al. 1985, Young and Hanson et al. 1987). *Pcf* was composed of 5' flanking and coding sequences of *atp9*, part of the *cox2* coding sequences and an unidentified open reading frame, *urfS* (Boeshore et al. 1985, Young and Hanson et al. 1987). Like the chimeric gene *T-urf13* in maize, the duplicated *atp9* and *cox2* were also present elsewhere in the petunia mitochondrial genome (Boeshore et al. 1985, Young and Hanson 1987). *Pcf* transcript patterns were similar between male-sterile lines and fertility restored line, thus, the *Rf* gene acted at the protein level, rather than at the transcriptional level (Nivison and Hanson 1989).

In radish, male-sterile Ogura cytoplasm has been investigated in detail. An open reading frame, *orf138* (Bonhomme et al. 1992, Krishnasamy and Makaroff 1993), generated from multiple rearrangements, was found 5' to *orfB*, a conventional mitochondrial gene (Quagliariello et al. 1990). *Orf138* is only found in Ogura radish, although sequences similar to *orf138* may be present in the normal radish (Krishnasamy and Makaroff 1993). Immediately 5' to *orf138* is a DNA sequence present 5' to normal radish *atp6* (Makaroff and Palmer 1988) and the Ogura *nad3/rps12*. Contrary to the situation in maize, the fertility restoration genes have no effect on *orf138* transcript patterns, but rather at protein level. ORF138 is produced in an organ-specific manner (Krishnasamy and Makaroff 1994).

In sorghum, several cms associated structures, including deletions/insertions within open reading frames, and chimeric configurations, have been described (Pring et al. 1995).

A novel observation in the sorghum chloroplast is that of a deletion of 55 residues of the *rpoC2* gene, coding for the ß-subunit of chloroplast RNA polymerase, was found to be correlated with male sterility in several male-sterile cytoplasms (Chen et al. 1993). An altered obligate gene product, COXI, of male-sterile cytoplasms of IS17218 (9E group) and IS7920C (A4 group) was also revealed to be associated with cms trait in sorghum (Bailey-Serres et al. 1986a, 1986b, Dixon and Leaver 1982). Mitochondria from the male-sterile 9E produced a 42 kD COXI product (Bailey-Serres et al. 1986a, 1986b), while mitochondria from other cytoplasms, including Milo cytoplasm (Pring et al. 1995), synthesized a 38 kD COXI polypeptide. A recombinational event at the 3' end of *cox1* generated an 303 base extended open reading frame for the 9E gene, relative to *cox1* in Milo (Bailey-Serres et al. 1986b). Two open reading frames, *orf265* (Tang et al. submitted) and *orf107* (Tang et al. in preparation) have been shown to be associated with cms in the male-sterile IS1112C cytoplasm of sorghum.

Other chimeric genes or genomic rearrangements, such as *orf224/atp6* in Polima *Brassica napus* (Singh and Brown 1991), *orf522* in sunflower (Köhler et al. 1991, Laver et al. 1991), *pvs* in the common bean *Phaseolus vulgaris* (Mackenzie and Chase 1990), and *orf79* in rice (Akagi et al. 1994), have also been demonstrated to be associated with the cms trait.

Transcription Initiation in Mitochondria

Transcription Initiation of Animal Mitochondrial DNA

Transcription of human mtDNA is initiated from two major promoters, the heavy strand promoter (HSP) and the light strand promoter (LSP) located in the D-loop region (Clayton 1991). Most of the genes are transcribed from the heavy (H)-strand, and virtually

the whole mitochondrial genome is transcribed. The DNA sequence and protein element requirements for specific initiation of transcription have been extensively characterized by utilizing an *in vitro* transcription system (Clayton 1991). Each of the two promoters, which spans about 50 nucleotides, requires two distinct sequence elements for activity. Analysis of deletion study (Chang and Clayton 1984), site-specific mutagenesis (Hixson et al. 1985), and linker-scanning mutagenesis (Topper and Clayton 1989) revealed that one element of 8-bp surrounding the transcription initiation site, 5'-CC(G)CCAAAA(+4)-3', is absolutely required for transcription activity. The second element is located upstream of the transcription initiation site and is required for high levels of specific initiation (Clayton 1991). Thus these two parts of DNA elements give rise the bipartite feature to human mitochondrial promoters (Jaehning 1993, Shadel and Clayton 1993). Human mitochondrial extracts have been partially purified into two active fractions: one containing a nonspecific mtRNA polymerase and the other containing a transcriptional factor (h-mtTFA), which is required for high levels of specific transcription initiation by binding to the upstream element of HSP and LSP (Clayton 1991).

The mouse LSP is similar to the human mtDNA promoters: A DNA sequence around the RNA start site is crucial for transcription, and the upstream sequences are required for efficient transcription (Chang et al. 1986a). Among the mitochondrial promoters analyzed to date, the HSP in mouse mtDNA appears to be unique. The DNA sequences required for transcription are located at positions -41 to -17, and the sequences surrounding transcription initiation site are not required for the specific initiation of transcription (Chang et al. 1986b).

Unlike the bipartite human mitochondrial promoters, the promoter sequences of *Xenopus laevis* mtDNA (Bogenhagen et al. 1988) contain a consensus octanucleotide

sequence, 5'-ACGTTATA-3', which is by itself sufficient for accurate transcription initiation.

The overall vertebrate mitochondrial transcription machinery is highly selective, and the vertebrate mtDNA promoters characterized thus far do not have significant sequence similarities with one another (Clayton 1991). The specificity can also be shown by the fact that in heterologous transcription assays, mitochondrial RNA polymerase from human cells does not initiate transcription at mouse mtDNA promoters and vice versa (Shadel and Clayton 1993). Consequently, mitochondrial promoter sequences, at least in vertebrate species, may be specific to individual or closely related species (Shadel and Clayton 1993).

Transcription Initiation of Yeast Mitochondrial DNA

The number of promoters of yeast mitochondrial genome have been determined by *in vitro* capping with guanylyltransferase and direct RNA sequencing (Christianson et al. 1983). Unlike the system in the human mitochondrial genome, with two closely located promoters (Clayton 1990), some 20 different primary transcripts were identified for yeast mtDNA (Christianson et al. 1983), and these promoters are spread throughout the genome. To search for DNA sequences that define the yeast mitochondrial promoter, sequences surrounding promoter regions were analyzed and a highly conserved nine-nucleotide sequence, 5'-ATATAAGTA(+1)-3', was obvious among all transcriptional start points (Christianson et al. 1983). Further analysis of deletion study and site-directed mutagenesis (Biswas et al. 1990) in an *in vitro* transcription system revealed the critical region required for promoter function to be certain nucleotides within the nine-nucleotide element. In general, the consensus sequence seemed to be the optimal promoter.

Yeast mtRNA polymerase is composed of a 145-kD core subunit encoded by the nuclear *RPO41* gene (Greenleaf et al. 1986), and a 45 kD specificity factor, mtTFB,

encoded by the nuclear *MTF1* gene (Costanzo and Fox 1990). Analysis of amino acid sequence of core subunit revealed that the *MTF1* gene is homologous to the bacteriophage T7 and T3 RNA polymerases (Costanzo and Fox 1990). Although the core alone can initiate transcription, specific factors are required for specific transcriptional activity (reviewed by Jaehning 1993). Mutational analysis of mtTFB, a 341-amino-acid protein, defined two regions that are important for normal function both *in vivo* and *in vitro* (Jang and Jaehning 1991). The predicted amino acid sequence of these two regions of mtTFB revealed limited similarity to the conserved region 2 and 3 of bacterial sigma factors (Jang and Jaehning 1991). These regions have been proposed to be involved in recognition of the bacterial promoter (Dombroski et al. 1992), DNA unwinding (Waldburger and Succkind 1994), and interactions between the sigma factors and the core RNA polymerase (Zhou et al. 1992). However, mutations in the sigma crucial promoter-contacting region did not eliminate the ability of mtTFB to initiate transcription specifically *in vitro* (Shadel and Clayton 1995), suggesting a different mechanism of promoter recognition of yeast RNA polymerase than that of bacterial RNA polymerases.

Transcription Initiation of Plant Mitochondrial DNA

The analysis of plant mitochondrial genomes is usually much more complicated than those of animal and yeast systems because the complex genome organization and frequent recombination events in plant mitochondrial genomes, which may raise the copy number of a gene or reshuffling of transcriptional regulatory elements (Gray et al. 1992). In the latter case, a homologous gene, in different plant mitochondrial genomes, may have diverged 5' or 3' flanking sequences, which may include transcriptional regulatory elements. A loose consensus sequence motif is expected for plant mtDNA promoters (Gray et al. 1992), compared with those highly conserved motifs for animal and yeast mtDNAs.

Transcript-mapping studies with primer extension, S1 nuclease mapping and *in vitro* capping with guanylyltransferase, transcription initiation sites of a number of maize mitochondrial genes, including *rrn18*, *rrn26*, *atp1*, *atp6*, *cox3*, *atp9*, *T-urf13*, and *orf221*, have been identified (Kennell and Pring 1989, McCarty et al. 1988, Mulligan et al. 1988a, 1988b, 1991). Multiple initiation sites have been detected, such as *cox2* and *atp9* with as many as five or six initiation sites (McCarty et al. 1988, Mulligan et al. 1991). A repeated segment 5' to *T-urf13* and *atp6* was found to drive transcription of both *T-urf13* and *atp6* (Kennell and Pring 1989).

In hope of identifying DNA motifs that define maize mitochondrial promoters, sequences surrounding the transcription initiation sites were compared for shared elements. A consensus sequence, (A/T)CRTA(G/T)A(A/T)AAA was revealed, and a sequence of CRTA is the most highly conserved feature of transcription initiation regions (Mulligan et al. 1991). In addition, the repeated motif G(A/T)$_{3-4}$ or C(A/T)$_{3-4}$ is found upstream of this element in some genes (Mulligan et al. 1991). However, DNA sequences of some initiation regions, such as five initiation sites of *atp9*, do not fit this proposed motif, even within the core CRTA region (Mulligan et al. 1991).

Similarly, several 5' transcript termini of wheat mitochondrial genes (*atp1/atp9*, *cox2*, *cox3*, and *orf25*) were also identified by isolating and directly sequencing capped transcripts (Covello et al. 1991). A consensus sequence, GCRTAtA, with the initiation site 2-3 nucleotides downstream of the final A, was defined (Covello et al. 1991).

Transcription initiation sites of a rice gene cluster, *rps3-rpl16-nad3-rps12*, were also determined (Nakazono et al. 1995a). Two initiation sites were identified: one was located between *rpl16* and *nad3*, and the other one was mapped upstream of *rps3* (Nakazono et al. 1995a). Alignment of the nucleotide sequences around the transcription initiation sites of

these genes, as well as other rice genes (*orf483*, *atp9* and *cob*), revealed a conserved CRTA element (Nakazono et al. 1995b).

Using similar transcript-mapping approaches, a number of 5' primary transcript termini of mitochondrial genes of dicots were mapped to the corresponding DNA sequences. In soybean, three mitochondrial gene, *atp9*, RNA *b/c* and RNA *e* (RNA b, c and e are three species of soybean mtRNA of approximately 150, 135 and 80 nt, respectively) transcription initiation sites have been identified (Auchincloss and Brown 1989, Brown et al. 1991). Alignment of initiation sites revealed that *atp9* and *b/c* transcripts shared the CRTA element, as proposed for maize system, whereas *e* transcript lacked this conserved motif (Brown et al. 1991).

In *Oenothera*, transcription initiation sites have been determined by *in vitro* capping of mtRNA from *Oenothera berteriana* (Binder and Brennicke 1993b). At least 15 initiation sites were identified, suggesting that the *Oenothera* mitochondrial genome contained a minimum number of 15 promoters (Binder and Brennicke, 1993b). Primary 5' transcript termini of *cox2*, *atp1* and an 18 S rRNA precursor have also been determined. Alignment of the sequences surrounding the 5' transcript termini revealed a conserved sequence block of 29 nt, aAaNTtgAaAtNTCaTAaGaGAagcaaaaN (Binder and Brennicke, 1993), with a 9-nt core (5'-CRTAaGaGA-3') conserved with soybean transcription initiation sites.

In potato, the transcription initiation site of *rrn26* was mapped (Binder et al. 1994). In contrast to the situation of *rrn18* in *Oenothera* (Binder and Brennicke, 1993b), the transcription initiation site of *rrn26* was determined to be the mature 5' end of *rrn26* transcript (Binder et al. 1994). Thus, in potato mitochondria, the mature 5' end of the *rrn26* is created by transcription initiation without any further 5' processing. The nucleotide sequence surrounding the initiation site shows only limited similarity to other putative promoter sequence from dicot plant mitochondria, and lacks the CRTA motif (Binder et al.

1994). The transcription initiation site of another potato gene, *atp9*, has also been identified, and the sequence alignment revealed that the *atp9* promoter does not contain the CRTA element (Lizama et al. 1994).

The transcription initiation site of a tRNA has also been investigated in *Oenothera* (Binder and Brennicke 1993a). A $tRNA^{Phe}$ gene transcription initiation site was determined by primer extension and *in vitro* capping experiments. Sequences covering the transcription start point contain the CRTA conserved block, suggesting that tRNA, rRNA and mRNA can be transcribed from homologous promoters (Binder and Brennicke 1993a).

The availability of *in vitro* transcription system, initially established in wheat mitochondria (Hanic-Joyce and Gray 1991), enables functional analysis of plant mitochondrial promoters. With this system, the *cox2* promoter of wheat was localized to a 372-bp region spanning positions -56 to -427, within the sequence GTATAGTAAGTA (initiation site underlined) (Hanic-Joyce and Gray 1991). A YRTA (TATA) motif is apparent immediately 5' to the initiation site.

In maize, Rapp and Stern (1992, Rapp et al. 1993) were able to dissect the maize *atp1* promoter with a partially fractionated maize mitochondrial extract. An 11-nucleotide element, (A/T)CRTA(G/T)A(A/T)AAA, surrounding the initiation site, was identified to be indeed critical for promoter function. However, this DNA block alone was insufficient to direct transcription initiation *in vitro*. In contrast, deletion or alteration of upstream G(A/T)3-4 repeats had no effect on transcription *in vitro*. Specific transcription initiation of maize *cox3* and *atp6* were also obtained in the *in vitro* system. The promoter of the *rrn18* gene, however, was not recognized by this RNA polymerase preparation, although the *rrn18* promoter bore a YRTA (corresponding to the CRTA proposed initially) core motif. These ramifications suggested that a more complex transcription machinery may required for faithful initiation of plant mtDNA.

Further analysis with linker-scanning and point mutagenesis (Rapp et al. 1993) showed that the *atp1* promoter consisted of a central domain extending from -7 to +5, and an upstream domain of 1 to 3 bp, centered around -11 to -12. The core region bearing YRTA corresponded to the conserved CRTA proposed by Mulligan et al (1991).

A promoter of perennial teosinte *cox2* gene was also defined in an *in vitro* transcription system with maize extract (Newton et al. 1995), and a conserved YRTA motif was found within the promoter region. Another *cox2* promoter, which is nuclear gene *Mct*-dependent and is only active when the dominant allele of *Mct* is present, was not able to initiate transcription in this *in vitro* maize system, although the *in vitro* extract was made from a maize line that carried the dominant *Mct* allele in the nucleus (Newton et al, 1995).

To date, nothing is known regarding the nuclear-encoded factors involved in the transcription initiation of plant mitochondrial genes. From studies of animal and yeast systems (Jaehning 1993, Shadel and Clayton 1993), plant mitochondrial RNA polymerase is expected to be encoded by the nuclear genome. Based on the similarity (YRTA) in promoter structure, it has been speculated that the polymerase from fungi, vertebrates and plants may have a similar basic structure of a core polymerase which is related to those utilized by bacteriophage T7 and T3, and an important specific factor for site selection (Jaehning 1993).

Regulation of Mitochondrial Gene Expression

Compared with the studies of nuclear gene expression, regulation of mitochondrial gene expression is poorly understood. Experiments of incorporating 5-bromodeoxyuridine into mtDNA of a mouse fibroblast cell line lacking the cytosolic thymidine kinase suggested that transcription of mtDNA is constitutive and not rate limiting for protein synthesis (Attardi

and Schatz 1988). Posttranscriptional regulation may play a major role in mammalian mitochondria, as exemplified by the differential stability of HeLa cell mitochondria. This differential stability contributes to the 20- to 150-fold higher copy number of individual tRNAs relative to the individual mRNAs (Attardi and Schatz 1988). Regulation may also occur at the translation level of mammalian mitochondrial genes, although the mechanism is not understood yet.

In yeast mitochondria, regulation of transcriptional initiation plays a major role in differential gene expression. As much as a 20-fold difference was revealed between the strongest and the weakest promoters (Costanzo and Fox 1990). Mitochondrial transcription may be modulated under different physiological conditions (Costanzo and Fox 1990). The steady state level of mitochondrial mRNAs in glucose-depressed cells were shown to be three- to six-fold higher than the mRNAs in repressed cells (Costanzo and Fox 1990). Yet, there is also evidence of no decrease in mtDNA transcription during anaerobic growth (Attardi and Schatz 1988).

Studies of differential gene expression in plant mitochondria have been conducted. Labeling of isolated mitochondria from organs of maize demonstrated that mitochondrial proteins synthesized in organelle differ among shoots, cobs, tassels and scutella of maize (Newton and Walbot 1985). Expression of a cms-associated gene, *orf138*, in fertility-restored Ogura radish was revealed as an organ-specific reduction at the protein level in flowers and leaves, but not roots (Krishnasamy and Makaroff 1994). Localization of two specific mitochondria-encoded proteins in petunia demonstrated that mitochondria in the same tissue could differentially express mitochondrially encoded proteins (Conley and Hanson 1994). The ATP1 protein was found at high levels in sporogeneous tissue and vascular bundle and at low levels in the connective and subepidermal tissues. In contrary, the COX2 protein was localized at high levels in the connective and subepidermal tissues but

at low levels in sporogeneous and vascular bundle (Conley and Hanson 1994). Apart from the regulation at the protein level, differences at the transcript level have also been studied. Transcripts and protein of *atp1* and cms-associated *orf522* in sunflower are abundant in young meiocytes of cells and tapetal cells, and were reduced in the presence of fertility restoration genes (Smart et al. 1994). Nevertheless, there is no evidence to date that differences at the transcript or protein levels are correlated with specific promoters.

MATERIALS AND METHODS

Genetic Stocks

Sorghum lines Tx398, IS1112C and A3Tx398 were provided by Dr. K. F. Schertz, USDA-ARS, College Station, Texas, and were maintained by Dr. D. R. Pring. Tx398 is a normal male-fertile line. IS1112C is a male-fertile line, carrying the male-sterile IS1112C cytoplasm. A3Tx398 is male-sterile line, carrying the IS1112C cytoplasm in the Tx398 nuclear background.

The maize hybrid line was provided by Pioneer Hybrid International, Johnston, Iowa.

Isolation of Mitochondrial DNA, RNA and Protein

Isolation of Mitochondria

Mitochondria were prepared from five to seven-day-old etiolated coleoptiles of sorghum or maize as previously described (Chase and Pring, 1986). Tissues were ground in a mortar and pestle or a Waring blender with 10 ml/gm tissue or 5 ml/gm tissue (for RNA isolation) of ice-cold buffer A [0.5 M sucrose, 5 mM Na_2EDTA, 0.1% BSA, 2 mM mercaptoethanol, 5 mM DIECA (for DNA isolation), 50 mM Tris, pH7.5], and filtered

through two layers of cheesecloth and two layers of miracloth. The filtrate was centrifuged at 1000 g for 10 minutes at 4C to remove cell debris and nuclei, and the supernatant was transferred to a clean tube and centrifuged at 15,900 g for 20 minutes at 4C to pellet mitochondria.

Isolation of mitochondrial DNA

Mitochondrial DNA was isolated using the procedure described by Chase and Pring (1986). After mitochondrial pellets were prepared, they were resuspended in 45 ml/10-30 gm tissue of ice-cold buffer G (0.3 M sucrose, 50 mM Tris, pH7.5) and centrifuged at 1000 g for 10 minutes at 4C. The supernatant was transferred to a clean tube and centrifuged at 15,900 g for 20 minutes at 4C. The pellet was resuspended in ice-cold buffer G with a small brush to 10 ml/50 gm tissue. Magnesium chloride and deoxynuclease were then added to the final concentration of 10 mM and 20 µg/ml, respectively, and the sample was placed at 4C for 60 minutes. At the end of incubation, the sample was diluted to 20 ml with ice-cold buffer G, underlaid with 20 ml of shelf buffer (0.6 M sucrose, 20 mM Na_2EDTA, 10 mM Tris, pH7.2) and centrifuged at 15,900 g for 20 minutes at 4C. The pellet was resuspended in shelf buffer to 10 ml/50-100 gm tissue, and centrifuged at 15,900 g for 10 minutes at 4C. The mitochondrial pellet was resuspended in 2.5 ml/50 gm tissue of lysis buffer (50 mM Tris, pH8.0, 20 mM Na_2EDTA). Sodium dodecyl sulfate and Proteinase K were added to 0.5% and 100 µg/ml, respectively. The sample was mixed gently and incubated at 37C for 60 minutes. At the end of the incubation, a 0.026 volume of 8 M ammonium acetate was added and the mixture was extracted twice with an equal volume of 75% water-saturated phenol, 25% chloroform-isoamyl alcohol (24:1) and once with equal volume chloroform-isoamyl alcohol (24:1) by emulsifying thoroughly and centrifuging at 7,000 g for 15 minutes. The upper

aqueous layer was collected and the DNA was precipitated with 0.1 volume of 8 M ammonium acetate and two volumes of absolute ethanol at -80C for 60 minutes. The pellet was collected by centrifuging at 15,900 g for 30 minutes at 4C, rinsed twice with absolute ethanol by centrifuging for 10 minutes, and drained dry. The DNA was resuspended in 200 µl/50 gm tissue of sterile distilled H_2O, and 10 µl of ribonuclease mixture (1.02 mg/ml RNase A, 10,000 U/ml RNase T1, boiled for 5 minutes) was added and incubated at 37C for 30 minutes.

Isolation of Mitochondrial RNA via Guanidine Thiocyanate and LiCl

Mitochondria were prepared as described above, and mitochondrial RNA (mtRNA) was isolated by the method of Kennell and Pring (1989). The mitochondrial pellet was resuspended in 10 ml buffer A, underlayed with 20 ml shelf buffer and centrifuged at 15,900 g for 20 minutes. The pellet was resuspended in 1-2 ml of buffer A and a 1.5 ml of guanidine thiocyanate lysis solution (6 M guanidine thiocyanate, 50 mM Tris, pH7.6, 10 mM Na_2EDTA and 2% N-Lauroyl sarcosine) was added to the sample. The lysate was dispersed with a brush, and 0.026 volume of 8 M ammonium acetate was added. The sample was extracted twice with equal volume of 75% water-saturated phenol, 25% chloroform-isoamyl alcohol (24:1) and once with equal volume of chloroform-isoamyl alcohol (24:1). The upper aqueous layer was collected, brought to 5 ml/100 gm tissue with buffer A, and a 0.2 volume of sterile 12 M LiCl was added, and incubated overnight at 0C. The sample was centrifuged at 15,900 g for 20 minutes at 4C, and the pellet was resuspended in 5 ml/100 gm tissue of sterile distilled H_2O. A 0.2 volume of sterile 12 M LiCl was added and incubated overnight at 0C. The sample was centrifuged at 15,900 g for 20 minutes and the pellet was resuspended in 200 µl/100 g tissue of sterile distilled H_2O.

Isolation of Mitochondrial RNA via Percoll Gradients

Mitochondria were prepared as described above except 2 mM mercaptoethanol was replaced with 4 mM dithiothreitol (DTT). The pellet was resuspended in ice-cold 0.75 X mitochondrial wash buffer (1 X: 0.25 M sucrose, 0.5 mM Na_2EDTA, 15% glycerol, 1 mM DTT, 20 mM Tris, pH7.5) to a final volume of 6 ml/100 gm tissue. Percoll gradients (Neuburger et al. 1982) were prepared with 45%, 21% and 13.5% Percoll solutions (in 0.5 M sucrose, 2 mM Na_2EDTA, 0.4% BSA, 2 mM DTT and 20 mM Tris, pH7.5) and kept at 4C. The mitochondria were carefully layered onto the Percoll gradients and centrifuged at 23,500 rpm in a Beckman SW27 rotor for 30 minutes at 4C. The materials above the 45-21% interface were removed, the mitochondria from the 45-21% interface were collected, and diluted with a total of 10 volumes of 0.75 X mitochondrial wash buffer and centrifuged at 15,900 g for 20 minutes at 4C. The pellet was collected, diluted with 5 volumes of 0.5 X mitochondrial wash buffer, and centrifuged at 15,900 g for 20 minutes. The mitochondrial pellet was collected, resuspended in 3-4 ml of 0.5 X mitochondrial wash buffer, transferred to a 1.5 ml Eppendorf tube, and centrifuged at 16,000 g for 5 minutes.

The mitochondrial pellet was resuspended in 2 ml GTC lysis buffer (4 M guanidium thiocyanate, 0.5% sarkosyl, 0.1 M mercaptoethanol, 25 mM sodium citrate, pH7.0) with a brush, and allowed to incubate at 25C for 5 minutes. A 0.1 volume of 2 M sodium acetate (pH4.0), 1 volume of water-saturated phenol and 0.2 volume of chloroform-isoamyl alcohol (49:1) were added, emulsified thoroughly, and placed on ice for 15 minutes. The sample was centrifuged at 7,000 g for 10 minutes. The upper aqueous layer was collected, and the RNA was precipitated with one volume of isopropanol at -80C for 60 minutes. The pellet was collecting by centrifuging at 15,900 g

for 20 minutes, rinsed twice with absolute ethanol, and drained dry. The dried RNA was resuspended in 200 µl sterile distilled H_2O /100 gm tissue.

Isolation of Mitochondria for Protein Purification

Mitochondria were prepared as described above except with modified buffer A (0.5 M sucrose, 50 mM 3-[N-Morpholino]propane-sulfonic acid (MOPS), 0.1% BSA, 4mM DTT, 10 mM EGTA, pH7.5) and modified Percoll solution (in 0.5M sucrose, 2 mM EGTA, 0.4% BSA, 2 mM DTT, 20 mM MOPS, pH7.5). After mitochondria were pelleted in a 1.5 ml Eppendorf tube, the pellet was quickly frozen in liquid nitrogen and placed at -80C.

Recombinant DNA Methods

Cloning

In this study, plasmids pUC19 and pBSSK(-) were utilized as vectors for cloning. The vector DNAs were digested with appropriate restriction enzymes, and the desired fragments were isolated by gel electrophoresis. One hundred nanograms of the vector DNA and a triple molar amount of foreign DNA were added to a ligation mixture of 66 mM Tris, pH7.5, 5 mM $MgCl_2$, 1 mM DTT, 1 mM ATP, and 1 unit of T_4 DNA ligase (5 unit for blunt end ligation) (Boehringer Mannheim Biochemicals). The ligation reactions were performed at 25C for 16 hours. The ligated DNAs were added to a tube containing 200 µl competent *Escherichia coli* cells DH5α (the procedure for making competent *E. coli* is shown below), mixed and stored on ice for 30 minutes. The tube was then transferred to 42C water bath for 40 seconds and chilled on ice for 5 minutes. An equal volume of LB medium (1% of bacto-tryptone, 0.5% bacto-yeast extract, and 1% of NaCl) was added to

the tube and the tube was placed at 37C for 30 minutes. The transformed competent cells were spread onto agar LB medium with 20 μg/ml of ampicillin and 800 μg of X-gal (5-bromo-4-chloro-3-indolyl-ß-D-galactoside), and incubated at 37C for 16 hours.

To make competent DH5α, bacteria were streaked from a frozen stock (stored at -80) onto a LB agar plate, and were incubated at 37C for 16 hours. A single colony was picked up from the plate and transferred into 50 ml of LB broth. The culture was incubated at 37C for about 3 hours with vigorous shaking. The concentration of bacterial cells was monitored by measuring the OD600 and did not exceed 10^8 cells/ml. The cells were then transferred aseptically to a sterile tube and chilled on ice for 10 minutes. The bacterial cells were recovered by centrifugation at 1000 g for 10 minutes at 4C. The medium was discarded and the pellet was resuspended in 10 ml of ice-cold 0.1 M $CaCl_2$ and stored on ice. The cells were recovered by centrifugation at 1000 g for 10 minutes at 4C, and the pellet was resuspended in 2 ml of ice-cold 0.1 M $CaCl_2$. At this stage, the cells were ready for used in transformation. The cells could also be dispensed into aliquots that could be frozen at -80C.

After transformation, bacterial colonies that contain the recombinant plasmid were identified by α-complementation (Sambrook et al. 1989) in which deletion mutants of the operator-proximal element of bacteria *lacZ* gene in the bacteria host were complemented by ß-galactosidase-negative mutants of plamids which contain the intact operator-proximal region. Due to the α-complementation, the Lac^+ bacteria produced blue colonies in the presence of X-gal (Sambrook et al. 1989). The recombinant plasmids, however, formed white colonies due to the insertion of foreign DNA and thus were incapable of α-complementation.

Recombinant Plasmids

The recombinant plasmid clones utilized in the investigation were listed in Table 1.

Preparation of Plasmids

By triton lysis method

Transformed cells were grown in 45 ml LB broth containing 0.1 g/l ampicillin at 37C with agitation for 16 hours. The cells were collected by centrifuging 10 minutes at 6000 g, and resuspended in 0.58 ml ST solution (25% sucrose, 50 mM Tris, pH8.0). Hen egg white lysozyme (0.195 ml of 10 mg/ml) was added to the tube, vortexed, and allowed to stand 5 minutes on ice, followed by adding 0.475 ml of 0.2 M Na_2EDTA, pH8. The sample was placed on ice for 5 minutes, and 1.25 ml Triton lysis solution (0.1% Triton X-100, 50 mM Tris, pH8.0, 6.25 mM Na_2EDTA) was added. The sample was rotated gently, and placed on ice for 15 minutes. The cells were pelleted by centrifuging at 15,900 g for 20 minutes at 4C. The supernatant was collected, and extracted with phenol and chloroform, precipitated and dried, as described above. The dried DNA was resuspended with 500 µl/50ml culture of sterile distilled H_2O, and a 10 µl RNase mixture (1.02 mg/ml RNase A, 10,000 U/ml RNase T1, boiled for 5 minutes) was added, and incubated at 37C for 30 minutes.

By boiling method

E. coli cells were grown overnight in TB media (terrific broth, per liter: 12 g Bacto tryptone, 24 g Bacto yeast extract, 4 ml glycerol, 17 m M KH_2PO_4, 72 mM K_2HPO_4) (Sambrook et al. 1989). On the following day, 1.5 ml saturated culture was transferred to a 1.5 ml Eppendorf tube and the cells were pelleted by spinning 30 seconds in a microfuge

at maximum speed. The supernatant was discarded and an additional 1.5 ml culture was added to the tube and the cells were pelleted. The pellet was resuspended thoroughly in a 400 µl ice-cold STET buffer (8% sucrose, 0.05% Triton X-100, 50 mM Na_2EDTA, 50 mM Tris, pH 8.0). Four hundred micrograms of lysozyme were added and the tube was held at 100C for 40 seconds. The sample was centrifuged for 15 minutes at 16,000 g, the supernatant collected, and the DNA was precipitated with 0.1 volume 8 M ammonium acetate and an equal volume of cold isopropanol at -80C for 60 minutes. The DNA was pelleted by centrifuging at 16,000 g for 20 minutes and rinsed twice with absolute ethanol. The near-dry DNA was resuspended in 50 µl sterile distilled H_2O.

DNA and RNA Analysis

Southern blotting

Mitochondrial DNA was digested with restriction endonucleases, and fractionated in agarose gels in 1 X TPE electrophoresis buffer (0.036 M Tris, 0.03 M NaH_2PO_4, 1 mM Na_2EDTA) with a constant voltage of 2 v/cm. After electrophoresis, the gel was placed in ethidium bromide staining solution (EB, 0.4 µg/ml) for 30 minutes, and photographed with transmitted UV illumination. The gel was treated with 0.25 M HCl for 10 minutes, rinsed with distilled water and denatured in 0.6 M NaCl and 0.2 M NaOH for 30 minutes. The gel was rinsed with distilled water and neutralized in 3 M NaCl and 1 M Tris, pH7.5, for 30 minutes. Schleicher & Schuell nitrocellulose membranes (BA-S 85) were prepared, wetted in distilled water, and placed in a tray of 20 X SSC (1 X: 0.15 M NaCl, 15 mM sodium citrate) (Sambrook et al. 1989) for 5 minutes. Three sheets of Whatman 3MM paper were placed on a 3-cm-thick stack of paper towels, and one piece of nitrocellulose membrane was laid on the top of the 3MM paper. After neutralization, the

gel was placed on the top of the nitrocellulose membrane and air bubbles were removed with a roller. Another nitrocellulose membrane was removed from the tray and laid on the top of the gel. Air bubbles were squeezed out by rolling a roller over the surface. Another three sheets of Whatman 3MM paper and a 3-cm-thick stack of paper towel were placed on the top of the filter. The transfer was allowed to proceed for three hours, and the DNA was fixed by baking the filter for 2 hours at 80C.

Preparation and labeling of probe DNA

The probe DNA was obtained as a cloned fragment which was purified from the vector by restriction digestion followed by recovery from an agarose gel and elution using Qiaex gel extraction kit (Qiagen Inc.) as per manufacturer's direction.

The DNA was labeled by random oligonucleotide priming using a random primed DNA labeling kit (Boehringer Mannheim Biochemicals). Typically, 500 ng of DNA was denatured and labeled with 10-50 µCi [α-^{32}P]dCTP as per manufacturer's direction. The un-incorporated deoxyribonucleoside triphosphates were removed by chromatography on a Sephadex G-50 column.

Hybridization of DNA blots

The immobilized DNA membrane was placed in a hybridization tube with hybridization solution [3 X SSC, 10 X Denhardt's solution (1 X: 0.02% Ficoll 400, 0.02% polyvinylpyrrolidone 360, 0.02% bovine serum albumin, 0.01% sodium dodecylsulfate), and 10 µg/ml of sheared salmon sperm DNA] (Chase and Pring 1986), and incubated for one hour, with rotation, at 65C in the hybridization oven. The labeled DNA was denatured for 10 minutes at 100C, placed on ice for 5 minutes, and added to the hybridization solution. The hybridization proceeded for 16 hours at 65C. After

hybridization, the membrane was washed in 3 X SSC for 5 minutes at 65C, and washed twice in 0.3 X SSC for 15 minutes each at 65C.

tRNAs

Mitochondria were isolated from etiolated coleoptiles and RNA was extracted as described above. tRNAs were isolated from the total RNA sample by holding in 2 M LiCl at 4C overnight (Parks et al. 1984). The sample was centrifuged at 15,900 g for 20 minutes at 4C and the supernatant was collected. The RNAs were precipitated from the supernatant by adding 0.1 volume of 8 M ammonium acetate and 2.5 volume of absolute ethanol and placed at -80C for one hour. tRNAs were pelleted by centrifuging at 15,900 g for 20 minutes, washed twice with absolute ethanol and dried under vacuum. The dried tRNAs were resuspended in sterile distilled H_2O.

About 50 µg of the resuspended tRNA was then treated with 50 units of RNase-free DNase (Boehringer Mannheim Biochemicals) in a reaction buffer of 40 mM Tris, pH7.5, 6 mM $MgCl_2$, 10mM NaCl and 10 mM DTT at 37C for 15 minutes. The sample was extracted twice with an equal volume of 75% water-saturated phenol, 25% chloroform and once with chloroform-isoamyl alcohol (24:1). The tRNAs were precipitated and dried as described above. The dried tRNAs were resuspended in sterile distilled H_2O.

The resuspended tRNAs were then heated at 55C for 5 minutes, placed on ice for 5 minutes, and treated with 10 units of calf intestine alkaline phosphatase (Boehringer Mannheim Biochemicals) in 50 mM Tris, pH9.8, 0.1 mM Na_2EDTA, at 50C for 60 minutes. The sample was extracted with phenol and chloroform, and dried as described above. The dried RNAs were resuspended in sterile distilled H_2O.

The 5' termini of tRNAs were labeled with 10 units of T_4 polynucleotide kinase (New England Biolabs) and [γ-^{32}P]ATP in 70 mM Tris, pH7.6, 10 mM $MgCl_2$, 5 mM

DTT, 40 μCi [γ-^{32}P]ATP (3000 Ci/mmol) and 80 units of human placentral RNase inhibitor (RNAsin, Promega Corporation). The product was extracted with phenol and chloroform, and dried as described above.

Polymerase chain reactions (PCR)

Polymerase chain reactions were performed with 100 ng of template DNA and 2 units of Vent DNA polymerase (New England Biolabs) plus 50 pmol of each specified primer in 100 μl of a solution of 10 mM KCl, 20 mM Tris, pH8.3, 10 mM $(NH_4)_2SO_4$, 200 μM each deoxynucleoside triphosphates (dNTP), and 0.1% Triton X-100. The reaction mixture were cycled in a thermocycler (Coy Tempcycler) for 30 cycles (94C 1 minute; 2 minutes under annealing conditions; and 72C, 3 minutes), followed by extension at 72C for 10 minutes. The products were treated twice with equal volume of chloroform-isoamylalcohol (24:1) and precipitated with 0.1 volume of 8 M ammonium acetate and 2 volumes of absolute ethanol.

For cloning, the PCR products were treated with Klenow DNA polymerase (Boehringer Mannheim Biochemicals) to generate blunt ends, extracted with phenol and chloroform, and dried as described above. The resuspended products were cloned into the *Sma*I site of Bluescript SK(-) (Stratagene) unless otherwise specified.

DNA sequencing

DNA sequencing was carried out by the dideoxynucleotide chain termination method (Sanger et al. 1977) with a DNA sequencing Kit (United States Biochemical). Approximately 1 μg of single-stranded DNA (M13) or 3 μg of double-stranded plasmid DNA was used. Double-stranded DNA was denatured in 0.2 M NaOH and 0.2 mM Na_2EDTA at 25C for 5 minutes, followed by neutralization with 0.1 volume of 2 M

ammonium acetate, and precipitated with 0.1 volume of 8 M ammonium acetate and 2 volumes of absolute ethanol. Annealing was performed by adding 2 µl of Sequenase reaction buffer and 1 µl of 2.5 pmol of the specific primer to the 7 µl of redissolved denatured plasmid DNA or 1 µl of 0.5 pmol of M13 specific primer, -40, and 5 µl M13 DNA with 2 µl H_2O. The mixture was heated to 65C for 2 minutes, and was placed at room temperature for 20 minutes period. The following reagents were added to the ice-cold annealed DNA mixture: 1 µl of 0.1 M DTT, 2 µl of labeling mixture, 10 µCi of [α-^{35}S]dATP and 2 µl of Sequenase version 2.0. The mixture was incubated for 5 minutes at 25C, and 3.5 µl of the reaction was transferred to each A, C, G, T tube containing 2.5 µl of dideoxynucleotide termination mixture prewarmed at 37C. The incubation was continued for another 5 minutes at 37C and the reaction was stopped by adding 4 µl of termination reaction.

Primer extension analysis

Gene-specific synthetic oligonucleotides (1 pmol) were labeled with 40 µCi [γ-^{32}P]ATP (3000 Ci/mmol) in a reaction containing 50 mM Tris, pH8.5, 10 mM $MgCl_2$, 5 mM DTT, and 10 unit of T4 polynucleotide kinase (Boehringer Mannheim Biochemicals) at 37C for 30 minutes. Ten micrograms of mtRNA and 0.1 pmol labeled primer were denatured by heating to 75C for 5 minutes, and annealed at 50C for 5 minutes. The extension was performed in a reaction of 10 mM Tris, pH8.5, 6 mM $MgCl_2$, 50 mM KCl, 1 mM DTT, 1 mM deoxynucleotide triphosphate mixture (dATP, dCTP, dGTP, and dTTP) and 12 units of AMV reverse transcriptase (Boehringer Mannheim Biochemicals) at 42C for 15 minutes. The reaction was terminated by 5 µl of formamide stop buffer, and the sample was boiled for 3 minutes prior to loading on a 6% denaturing polyacrylamide gel.

Synthetic nucleotides used in primer extension analysis were listed in Table 2.

Characterization of Initiated Transcript Termini

In vitro Capping Reaction

One hundred micrograms of sorghum mtRNA extracted from mitochondria purified through Percoll gradients were capped in 50 mM Tris, pH7.9, 1.25 mM $MgCl_2$, 2.5 mM DTT, 6 mM KCl, 80 units of RNase inhibitor (RNasin, Promega Corporation), 1 μM GTP, 250 μCi [α-^{32}P]GTP (3000 Ci/mmol) and 12 units of guanylyltransferase (GIBCO BRL Life Technologies) at 37C for 30 minutes. The reaction was terminated with 20 μg of Proteinase K in 100 μl 5 mM Na_2EDTA and 0.5% SDS at 37C for 15 minutes. The capped RNA was extracted, precipitated and resuspend in sterile distilled H_2O, as described above.

Antisense RNA

Linearized templates in pBluescript vector were used for the respective genes to transcribe RNA *in vitro*. The transcription reaction was performed in 20 μl of 40 mM Tris, pH7.5, 6 mM $MgCl_2$, 2 mM spermidine, 10 mM NaCl, 10 mM DTT, 20 units of RNase inhibitor (RNasin, Promega Corporation), 0.5 mM NTP mixture (ATP, CTP, GTP, UTP), 200-500 ng template DNA and 15 units of T_3 or T_7 RNA polymerase at 37C for 30 minutes. The template DNA was removed by digestion with 50 units of RNase-free DNase (Boehringer Mannheim Biochemicals) at 37C for 15 minutes. The reaction was extracted, precipitated and resuspend in sterile distilled H_2O.

Ribonuclease Protection

Approximately 10 µg of capped sorghum mtRNA and 2 µg of antisense RNA were denatured at 85C for 5 minutes in 30 µl of hybridization solution of 80% deionized formamide, 40 mM PIPES (Piperazine-N, N'-bis[2-ethane-sulfonic acid]), pH6.4, 0.4 M sodium acetate, 1 mM Na_2EDTA, and hybridization was allowed to proceed at 37C for 16 hours. Three hundred ul of RNase mixture was added which contained 20 mM sodium acetate, 10 mM Tris, pH7.4, 5 mM Na_2EDTA, one unit of RNase-One (Promega Corporation) unless specified, and incubated at 20C for 30 minutes. The reaction were terminated by 3.3 µl of 10% SDS and precipitated with 0.1 volume of 8 M ammonium acetate and 2.5 volumes of absolute ethanol. The RNA was denatured by heating at 70C for 10 minutes before loading on a 6% polyacrylamide gel.

Sizing of the Protected Fragments

The size of the protected fragment of capped RNA was approximated by electrophoresis on 6% polyacrylamide gel, using M13 or mtDNA genomic clones of known sequences as markers. Due to the mass difference between a nucleotide chain and the deoxynucleotide marker chain, corrections were made by calculating the actual mass of the expected RNA product and DNA marker utilized in the comparison. The protected fragment was thus approximately sized, and the assignment of precise 5' termini was based on primer extension experiments.

RNA Polymerase Assay

Purification of RNA Polymerase from Maize

Mitochondria of Pioneer hybrid maize were isolated as described above for protein purification. The transcript competent fraction was isolated using the procedure described by Rapp and Stern (1992).

Mitochondria were resuspended at 4C in 5 ml of mitochondrial lysis buffer of 20 mM Tris, pH8.0, 0.2 mM Na_2EDTA, 1 mM DTT, 1 mM Pefabloc (Boehringer Mannheim Biochemicals) and 15% glycerol, and transferred to a sterile glass tissue homogenizer. Sterile distilled H_2O (2.12 ml) was added, followed by 20 stokes with the pestle. Triton X-100 (20%, 0.25 ml) was added, followed by 20 strokes with the pestle. Finally, 2.62 ml of 4 M KCl was added, followed by 20 strokes with the pestle. The mitochondrial lysate was placed on ice for 15 minutes with vortexing every 5 minutes. The lysate was centrifuged at 4C at 32,500 rpm (Beckman SW50.1, 100,000 g) for 1 hour and the supernatant was collected. Centrifugation of the lysate at 100,000 g yielded a membrane-free soluble protein fraction.

An equal volume of S-100 dilution buffer (90 mM Tris, pH8.0, 0.1 mM Na_2EDTA, 0.5 mM DTT, 2.5% glycerol) was added to the S-100 supernatant, and the fraction was subjected to fractional precipitation with 20% ammonium sulfate by centrifuging at 12,000 rpm for 15 minutes with Sorvall SS34 rotor. The supernatant was further subjected to 20-50% ammonium sulfate precipitation and the protein was precipitated by centrifuging at 12,000 rpm for 1 hour with Sorvall SS34 rotor. The pellet was resuspended in 2 ml of buffer of 10 mM Tris, pH8.0, 1 mM DTT, 0.1 mM Na_2EDTA, 7.5% glycerol, 50 mM of KCl, and dialyzed against two changes of the same buffer.

The dialyzed 20-50% ammonium sulfate cut was loaded onto a mono-Q FPLC column (Pharmacia LKB Biotechnology) and was eluted by step-wise increases in KCl concentration of 0.1, 0.2, 0.3, and 0.5 M (in the above buffer) with a elution rate 0.5 ml/minute. The protein eluted in 0.3 M KCl was collected, combined if within the same peak, and dialyzed in buffer A with 10 mM KCl. The dialyzed protein was concentrated to 100-150 µl using centricon-10 unit (Amicon). The concentrated protein was aliquoted, quickly frozen in liquid nitrogen, and stored at -80C. The concentration of the proteins was determined by the microassay of the BioRad measuring the absorbance at 595 nm.

In vitro Transcription Assay

In vitro transcription assays were performed in a 12.5 µl reaction of 10 mM Tris, pH7.9, 10 mM $MgCl_2$, 1 mM DTT, 0.5 mM each ATP, CTP and GTP, 25 µM UTP, 10 µCi [α-^{32}P]UTP (3000 Ci/mmol), 40 units of RNase inhibitor (RNAsin, Promega Corporation), 60 ng of linearized template DNA or 40 ng of insert DNA, and 20-40 µg of protein. The reaction was incubated at 30C for 30 minutes, and terminated by the addition of 37.5 µl of stop solution containing 4.8 M urea, 0.4 M sodium acetate, 5.3 mM aurintricarboxylic acid, 26 µg/ml tRNA and 0.8% SDS. The mixture was immediately extracted twice with equal volume of 75% water-saturated phenol and 25% chloroform and once with chloroform-isoamylalcohol (24:1). The reaction products were precipitated with 0.1 volume of 8 M ammonium acetate and 2.5 volumes of absolute ethanol, and washed twice with absolute ethanol. The dried products were resuspended in 80% formamide, 1 X TBE and electrophoresed in denaturing polyacrylamide gels with 1 X TBE electrophoresis buffer (0.089 M Tris-borate, 0.089 M boric acid and 2 mM Na_2EDTA, pH8.0) (Sambrook et al. 1989).

Table 1. Recombinant DNA clones used in the investigation.

Clone	Vector	Kb	Fragment	Features	References
pAF1	pUC18	6.5	BamHI	Includes *atp6-1* IS1112C	(1)
pJM37	pUC18	0.217	XbaI/EcoRI	Includes part of *atp6-1*	(2)
pBY144	pBSSK(-)	0.268	HindIII	Includes part of *atp6-1* T3=antisense	this work
pBY150	pBSSK(-)	0.262	DP19/DP46	5' to *atp6-1* T3-DP46 proximal T3=antisense	this work
pAF2	pUC18	5.0	BamHI	Includes *atp6-2* IS1112C	(1)
pBY148	pBSSK(-)	0.443	DP31/DP17	5' to *atp6-2* T3-DP17 proximal T3=antisense	this work
pRS34	pUC119	11	BamHI	Includes *atp9* Tx398	(3, 4)
pRS35	pUC119	1.4	PstI	Includes *atp9*	(3, 4)
pRS61	pUC119	6.8	SmaI	Includes 5' to & part of *atp9* IS1112C	(3, 4)
pRS62	pUC119	1.236	EcoRI	Includes *trn*fM & *atp9* Tx398	(3, 4)
pRS66	pUC119	0.785	PstI/EcoRI	Includes *atp9*	(3, 4)
pRS72	pUC119	0.457	EcoRI/PstI	Includes *trn*fM	(3, 4)
pBY132	pBSSK(-)	0.807	1201/DP48	5' to *atp9* T3-1201 proximal T3=sense	this work
pBY136	pBSSK(-)	0.524	DP47/1211	Includes *atp9* T3-1211 proximal T3=antisense	this work
pHC86	pUC19	1.223	BamHI	Includes *orf107* IS1112C	(6)
pHC103	pUC19	0.214	BamHI/XbaI	5' to *orf107*	(6)
pBY145	pBSSK(-)	0.214	BamHI/XbaI	5' to *orf107*	this work
pHT130	pUC19	0.378	AluI/AluI	Includes part of *orf107*	(6)
pBY124	pUC19	2.9	BglII	Includes *orf107* IS1112C	(6)
pBY128	pUC19	1.8	EcoRI/BamHI	5' to *orf107*	this work
pBY142	pBSSK(-)	1.8	EcoRI/BamHI	5' to *orf107*	this work
pHT98	pUC19	1.3	SphII/HindIII	Includes 5' to & part of *urf209*	(5)
pBY146	pBSSK(-)	1.0	BamHI	Includes 5' to & part of *urf209*	this work
pHT70	pUC19	5.8	BamHI	Includes 5' to & part of *orf265*	(5)

Table 1. Recombinant DNA clones used in the investigation (continued).

Clone	Vector	Kb	Fragment	Features	References
pHT75	pUC19	0.91	BamHI/KpnI	Includes 5' to & part of *orf265*	(5)
pBY133	pBSSK(-)	0.319	1211/DP92	5' to *orf265* T3-DP92 proximal T3=antisense	this work
pHT100	pUC19	0.8	EcoRI/KpnI	5' to *orf265*	(5)
pBY149	pBSSK(-)	0.8	EcoRI/KpnI	5' to *orf265*	this work

1. Kempken et al. 1991.
2. Mullen et al. 1992.
3. Salazar et al. 1991.
4. Salazar 1991.
5. Tang et al. submitted.
6. Tang et al. in preparation.

Table 2. Synthetic oligonucleotides utilized in primer extension analyses. All the oligonucleotides listed were synthesized by the DNA synthesis laboratory, Interdisciplinary Center for Biotechnology Research, University of Florida.

Gene/orf	Oligonucleotide	Sequence	bp 5' to start codon (5' end of the primer)
atp6-1	DP13	5'-GTGACGTGGGTTTTACCCAGCG-3'	-52
	DP46	5'-CCACTTAGAAATATAAATAAAAAG-3'	-241
atp6-2	DP63	5'-GGGCGTTGGTGTTTCCAACGA-3'	-59
	DP49	5'-AAGACAAAATGCTAGCTTCCC-3'	-116
atp9	DP48	5'-CCTTCCTCCTCTCCTTAGTTG-3'	-68
	DP73	5'-GGTGATCAGGACTCGTACAAG-3'	-288
orf107	DP102	5'-TCCCATTCCAAATACCGGCGG-3'	+65
	DP103	5'-ATTTGTCTGGGATTGCGGCC-3'	-364
orf265	DP127	5'-GAACACGGACTGCCATTAC-3'	+296
	DP06	5'-AGACTCGATTCTCCCGTCGAGAG-3'	+131
	DP131	5'-GTAGGGAGCACACAACTTCTC-3'	-31

RESULTS

The transcript patterns of the six selected sorghum mitochondrial genes and open reading frames (orfs) utilized in this study have been described (Mullen et al. 1992, Salazar 1991, Tang et al. submitted, Tang et al. in preparation). Thus, to map the transcription initiation sites corresponding to the 5' primary transcript termini of selected genes, gene-specific fragments which include 5' flanking sequence of the gene were cloned and sequenced, and antisense RNA was synthesized *in vitro*. Since 5' termini of primary transcripts carry tri- or di-phosphates (Christianson et al. 1983), while processed transcripts have only monophosphate or hydroxyl groups at their 5' ends, primary transcripts could be labeled by capping at their 5' ends with the enzyme guanylyltransferase. *In vitro* capping in combination with ribonuclease protection was conducted to identify the 5' termini of the primary transcripts. Total mitochondrial RNA was prepared from mitochondria purified through Percoll gradients, and used in an *in vitro* capping reaction with α-^{32}P-GTP and guanylyltransferase. The capped RNA was hybridized to the antisense RNA synthesized from the gene-specific clones, followed by ribonuclease treatment. The protected fragment was analyzed to give an approximate size. To precisely map the 5' transcript ends, primer extensions with gene specific primers were performed.

In this study, the 5' transcript termini of six sorghum mitochondrial genes and open reading frames, including two orfs which are associated with CMS in sorghum, have been identified, and the transcription initiation sites have been precisely positioned. The

sequences surrounding initiation sites were compared and aligned. The promoter of sorghum mitochondrial *atp6-2*, which is homologous to that of maize *atp6* of mitochondrial DNA, was tested in an *in vitro* maize transcription system, and the transcription initiated at a correct site.

Transcription Initiation of *atp6-1*

Sorghum mitochondrial *atp6* occurs as one or two copies, depending on cytoplasm (Mullen et al. 1992). All sorghum lines examined to date carry a copy designated as *atp6-2* (Mullen et al. 1992, Pring et al. 1988,). This form of the gene consists of an open reading frame of 1137 bp, including a 756 bp conserved core region as those in yeast and other higher plant systems, flanked 5' by a 381 bp amino extension (Kempken et al. 1991, Mullen et al. 1992). The mature ATP6 polypeptide, from the core region, probably results from proteolytic cleavage, generating uniform amino termini (Krishnasamy et al. 1994).

An additional copy of the *atp6* gene, *atp6-1*, is found in type members of the A1 (three dwarf white milo), A2 (IS12662C) and A3 (IS1112C) groups of male-sterile cytoplasms of sorghum (Mullen et al. 1992, Pring et al. 1988). This version consists of an open reading frame of 1002 bp, sharing an identical core region of 756 bp as that in *atp6-2*, and flanked 5' by divergent extension of 246 bp (Mullen et al. 1992). A schematic diagram of the core regions, amino extensions of both copies is shown in Figure 1. Sequences 5' to *atp6-1* and *atp6-2* exhibit limited sequence similarity for about 100 bp, and then diverge, which indicates distinct 5' transcript termini for the *atp6-1* and *atp6-2* transcripts (Pring et al. 1992).

Both versions of *atp6* are transcribed (Mullen et al. 1992), and their transcripts are edited (Kempken et al. 1991). The editing events include 18 C-to-U substitutions within

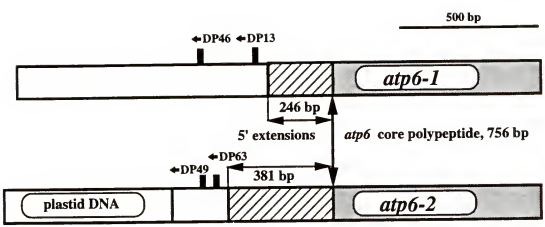


Figure 1. Schematic diagram of sorghum mitochondrial *atp6-1* and *atp6-2* core regions (dotted box), amino extensions (striped box) and the locations of the primers (small black box) utilized in primer extension experiments.

the core polypeptide and one edit at the carboxy extension, altering 15 amino acids and resulting in a carboxyl-terminus identical to that of yeast (Kempken et al. 1991). The amino extension of *atp6-2* has one silent edit (Kempken et al. 1991), while the amino extension of *atp6-1* exhibits one edit that changes one amino acid (Mullen et al. 1992).

Transcription of *atp6-1* is characterized with a single transcript of 1500 nt (Mullen et al. 1992). To map the transcription initiation site of *atp6-1*, the location of the 5' terminus of the primary *atp6-1* transcript was identified by ribonuclease protection with a gene-specific fragment.

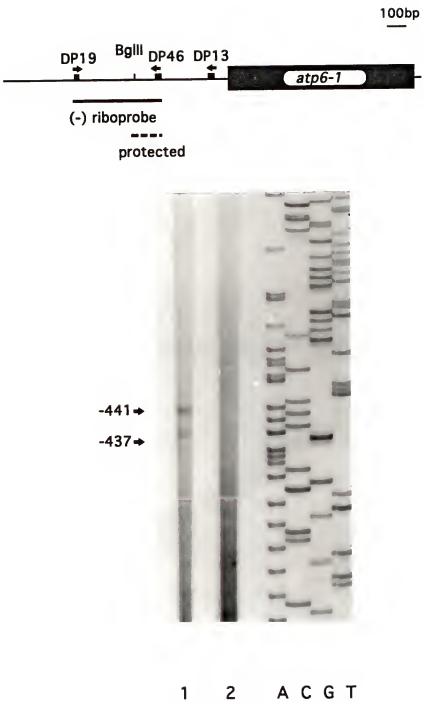
Cloning Gene Specific Fragment

A region 5' to *atp6-1* was cloned. Oligonucleotide DP19, 5'-GCTCATATCATTCTGCGTGC-3', corresponding to a sequence between -587 and -565, and oligonucleotide DP46 complementary to a sequence between -343 and -320 5' to the start codon (Table 2), were used in polymerase chain reaction with pAF1 DNA as a template. The amplified fragments were cloned into *Sma*I-digested pBSSK(-). Sequence analysis revealed that clone pBY150 had the DP46 end adjacent to the T3 promoter, and that the T3 RNA polymerase initiated antisense strand transcription of the cloned fragment.

In vitro Capping and Ribonuclease Protection

In vitro capping was carried out to identify the 5' primary transcript end(s) of *atp6-1*. Bluescript clone pBY150 DNA was used as a template for synthesis of antisense RNA for ribonuclease protection experiments. Single-stranded antisense RNA was transcribed *in vitro* from *Eco*RI-digested pBY150 DNA with T3 RNA polymerase, and hybridized to the capped RNA, followed by ribonuclease digestion. The protected fragment was analyzed as shown in Figure 2. The 5' termini of the protected product were assigned to

Figure 2. Ribonuclease protection of the *atp6-1* transcript. Antisense RNA (solid line) and sense RNA synthesized from Bluescript clone pBY150 were hybridized with capped mtRNAs, respectively, and treated with ribonuclease. The protected region (dashed line) was resolved with M13 sequence as a marker. Lane 1: Capped RNAs were hybridized with antisense RNA. Lane 2: Capped RNAs were hybridized with sense RNA.


100bp
DP19
BglII
DP46
DP13
atp6-1
(-) riboprobe
protected
-441
-437
1
2
A
C
G
T

the corresponding DNA sequence of about -441 and -437 nucleotide 5' to the start codon. As a control, single stranded sense RNA was also synthesized with T7 RNA polymerase using *Bam*HI digested pBY150 DNA as a template. As expected, no protected signal was detected (Figure 2).

Primer Extension

To precisely determine the 5' end(s) of the *atp6-1* transcript, oligonucleotide DP13 (Table 2), complementary to a sequence between -51 to -72 5' to the start codon, was prepared. Mitochondrial RNA from line IS1112C was used in a primer extension reaction with primer DP13, and the extended product was analyzed as shown in Figure 3A. A major product of about 392 bp was identified, with a terminus corresponding to the nucleotide of -442 5' to the start codon (Figure 4).

Similarly, another primer, DP46 (Table 2), which is 270 bp upstream of DP13, complementary to a sequence between -241 to -264 was utilized in primer extension experiments. A major extended product of 122 bp was detected (Figure 3B), with the end corresponding to nucleotide -442 (Figure 4). In addition, a minor product of 119 bp was also detected, with the terminus corresponding to nucleotide -439 5' to the start codon.

Transcription Initiation of *atp6-2*

Transcription of this copy is characterized by a major transcript of 1450 nt (Mullen et al. 1992). A schematic diagram of *atp6-2* is shown (Figure 5), including locations of the primers that were used in the following analyses.

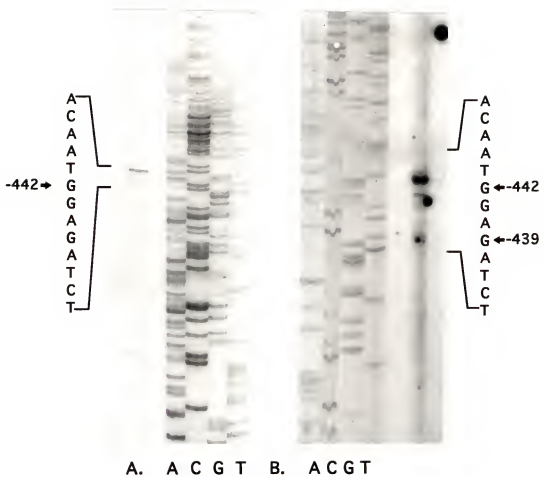


Figure 3. Primer extension analysis of *atp6-1*. A: Oligonucleotide DP13 was used in dideoxy sequencing of pAF1 and the primer extension reaction. B: Oligonucleotide DP46 was used in dideoxy sequencing of clone pBY150 and the primer extension reaction with sorghum mtRNA.

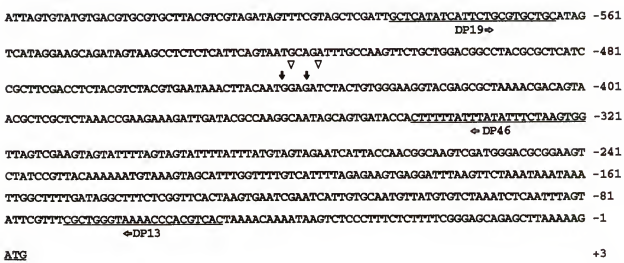
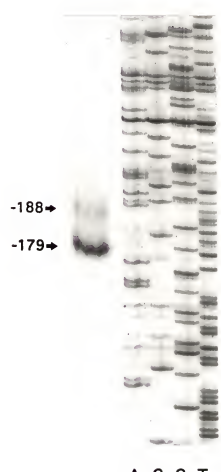


Figure 4. Nucleotide sequence of the 5' flanking region of *atp6-1*. The closed arrows represent the major termini mapped by primer extension experiments. The locations of the oligonucleotides are underlined. The open triangles represent the termini of the protected fragments from *in vitro* capping in conjunction with ribonuclease protection experiments.

Figure 5. Ribonuclease protection of the *atp6-2* transcript. Antisense RNA (solid line) synthesized from Bluescript clone pBY148 was hybridized with capped mtRNA, and treated with ribonuclease. The protected region (dashed line) was analyzed with clone pBY148 sequence as a marker.

DP31
DP17
100bp
BglII
atp6-2
(-) riboprobe
protected


-188
-179
A C G T

Cloning Gene Specific Fragment

A region 5' to the *atp6-2* gene, corresponding to a sequence between +105 to -338 5' to the start codon, was cloned. Oligonucleotide DP31, 5'-GCATAGATCCGGAGATTCCC-3', corresponding to a sequence between -338 and -317, and oligonucleotide DP17, 5'-CACCGGGGGCATCG-3', complementary to a sequence between +92 and +105, relative to the start codon, were prepared and used in the polymerase chain reaction with pAF2 DNA (Mullen et al. 1992) as a template. The amplified fragments were cloned into *Sma*I digested pBSSK(-). Sequencing analysis revealed that clone pBY148 had the DP17 end adjacent to the T3 promoter, and that the T3 RNA polymerase initiated antisense strand transcription of the cloned fragment.

In vitro Capping and Ribonuclease Protection:

Bluescript clone pBY148 was used as a template for transcribing antisense RNA *in vitro*. Single stranded antisense RNA was synthesized from *Bam*HI digested pBY148 DNA with T3 RNA polymerase, and hybridized to the capped RNA, followed by ribonuclease digestion. The protected fragment was analyzed (Figure 5). The 5' terminus of the protected product was assigned to the corresponding DNA sequence of about -179 5' to the start codon. Another minor 9 nt longer protected fragment was also detected, with a 5' terminus corresponding to a residue of -188.

Primer Extension Analysis:

To precisely map the 5' end(s) of the *atp6-2* transcript, oligonucleotide DP63 (Table 2), complementary to a sequence of -59 to -79, relative to the start codon, was prepared. Mitochondrial RNA was analyzed with primer DP63 in primer extension reactions. A

major extended product of 131 bp was identified (Figure 6A) with terminus corresponding to -188 5' to the start codon.

To verify the result obtained with primer DP63, another primer, DP49 (Table 2), complementary to a gene specific sequence of -116 to -136, was used in primer extension experiment. A major extended product of 73 bp was detected (Figure 6B), with terminus corresponding to nucleotide -188, relative to the start codon (Figure 7), thus confirming the result obtained with primer DP63.

Transcription Initiation of *atp9*

Sorghum mitochondrial *atp9* occurs as one copy in normal line Tx398, and two identical copies in some lines, such as IS1112C, A3Tx398 and A3Tx7000 (Salazar 1991, Salazar et al. 1991). The coding region of three identical versions consists of an open reading frame of 237 bp and would encode a 79 amino acid polypeptide (Salazar et al. 1991).

Transcription of sorghum *atp9* is characterized by a single transcript of 650 nt in all the lines examined (Salazar 1991). A total of eight C-to-U editing events were revealed within the coding region of the copy in either line Tx398 or IS1112C, changing six amino acids and generating a new stop codon which eliminates 12 residues predicted from the genomic sequence (Salazar et al. 1991).

Sequence analysis showed *trn*fM is located -323 bp 5' to all three copies of *atp9* (Salazar 1991), which is 98.7% identical to *trn*fM of maize (Parks et al. 1984) and wheat (Sprinzl et al. 1989). To identify the transcriptional properties of *atp9* and the transcription initiation site(s), *in vitro* capping experiments were conducted.

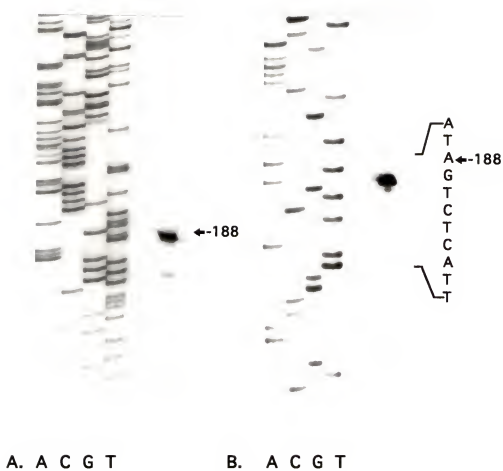


Figure 6. Primer extension analysis of *atp6-2*. A: Oligonucleotide DP63 was used for the primer extension reaction with M13 sequence as a marker. B: Oligonucleotide DP49 was used in dideoxy sequencing of clone pBY148 and the primer extension reaction with sorghum mtRNA.

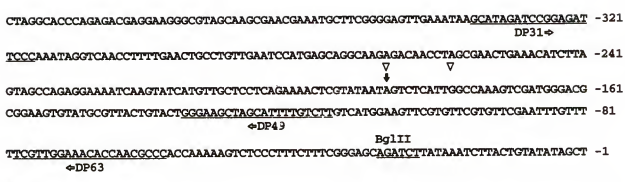


Figure 7. Nucleotide sequence of the 5' flanking region of *atp6-2*. The closed arrow represents the major terminus mapped by primer extension analysis. The locations of the oligonucleotides are underlined. The open triangle represents the terminus of protected fragment from *in vitro* capping in conjunction with ribonuclease protection experiments.

In vitro Capping

Hybridization of *in vitro* capped mRNA to a slot blot containing clones of the *trn*fM (Figure 8, A and B) and *atp9* (Figure 8, A and C) is shown. *In vitro* capped RNA strongly hybridized to subclones pRS62 (A) and pRS35 (C) and only poor hybridization was detected with subclone pRS72 (B). Since DNAs were loaded in equal amounts and only the primary transcripts were labeled (capped RNA), the observation that clones pRS62 and pRS35 gave much stronger hybridizations suggests that *atp9* and *trn*fM may be transcribed in different transcriptional units, or *atp9* may have its own initiation region in addition to sharing a precursor with *trn*fM. The second possibility was strongly supported by the results of ribonuclease protection.

Cloning Gene Specific Fragments for Ribonuclease Protection

Oligonucleotide 1201, 5'-AACAGCTAGACCATG-3', corresponding to a sequence of the vector of pRS62, and oligonucleotide DP48 (Table 2), complementary to a sequence between -68 and -88 relative to the start codon, were used in the polymerase chain reactions with pRS62 DNA (Salazar et al. 1991) as a template. The amplified fragments were cloned into *Sma*I digested pBSSK(-). Sequencing revealed that clone pBY132 had the 1201 end adjacent to the T3 promoter, and that the T3 RNA polymerase initiated sense strand transcription of the cloned fragment.

Oligonucleotide DP47, 5'-CAACTAAGGAGAGGAGGAAGG-3', corresponding to a sequence between -88 to -68, and oligonucleotide 1211, 5'-GTAAAACGACGGCCAGT-3', complementary to a sequence of pRS62 vector, were used in the polymerase chain reactions with pRS62 DNA as a template. The amplified fragments were cloned into *Sma*I digested pBSSK(-). Sequencing analysis revealed that

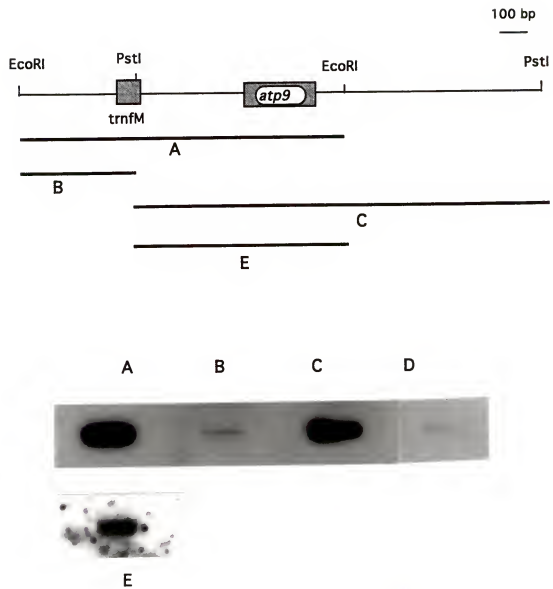


Figure 8. Hybridization of mtRNA capped with guanylyltransferase and [α-32P]GTP to a slot blot containing plasmid DNA of pRS62 (A), pRS72 (B), pRS35 (C) and vector pUC119 (D). In (E), a DNA blot containing clone PRS66 was hybridized to kinased mtRNA with polynucleotide kinase and [γ-32P]ATP.

the clone pBY136 had the 1211 end adjacent to the T3 promoter, and that the T3 RNA polymerase initiated antisense strand transcription of the cloned fragment.

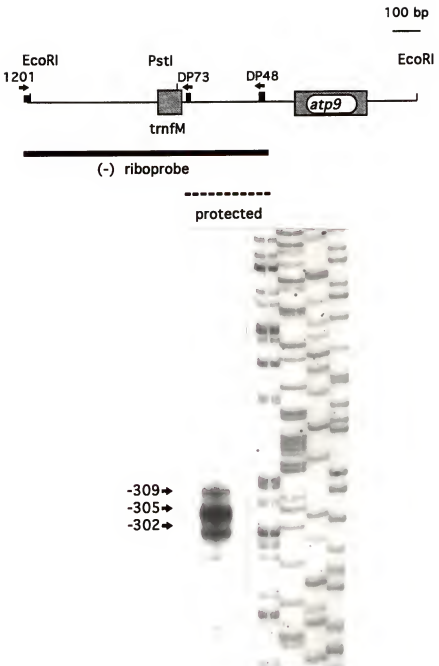
Ribonuclease Protection

Single stranded antisense RNAs were transcribed *in vitro* from clones pBY132 and pBY136, and hybridized with capped mtRNA, respectively, followed by ribonuclease digestion. The protected fragments were analyzed. Antisense RNA derived from clone pBY132 was used in the protection experiment (Figure 9), and the major protected fragments were sized as 241, 238 and 234 nt, with termini corresponding DNA sequence of -309, -305 and -302, relative to the start codon, which are located 3' to the 3' terminus of *trn*fM. This suggests that there is an initiation region 3' to *trn*fM that promotes transcription of *atp9*. Antisense RNA transcribed the downstream fragment pBY136 was also used in the protection experiment, and no signal was detected (not shown), suggesting that the pBY136 region lacked initiation termini. As controls, capped RNA with and without ribonuclease digestion were also analyzed.

Primer Extension Analysis

To precisely localize the 5' termini of the *atp9* transcript, oligonucleotide DP48 (Table 2), complementary to gene-specific sequences of *atp9*, was used in primer extension analyses with total mtRNA from line Tx398. Major extension products corresponding to -320, -319, -318, -314, -313, -312, -310, -306, -302, -296, -295, -291, -289, -279, -277, -275, -273, and -271, 5' to the start codon, were detected (Figure 10). The longest extended fragment ended at -322, the 3' terminus of *trn*fM (Figure 11). This pattern of multiple extension products was reproducible, and the RNAs used in these assays gave discrete termini in primer extension experiments with other genes. This

Figure 9. Ribonuclease protection of the *atp9* transcript. Antisense RNA (solid bar) synthesized from Bluescript clone pBY132 was hybridized with capped mtRNA, and treated with ribonuclease. The protected region (dashed line) was analyzed with M13 sequence as a marker.


100 bp
EcoRI
1201
PstI
DP73
DP48
EcoRI
atp9
trnfM
(-) riboprobe
protected
-309
-305
-302
A C G T

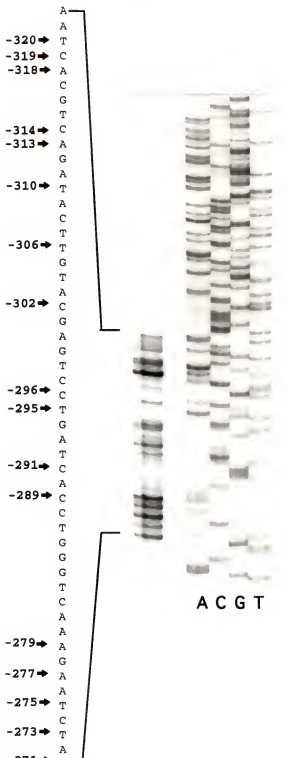


Figure 10. Primer extension analysis of *atp9*. Oligonucleotide DP48 was used in dideoxy sequencing of pBY132 and primer extension reaction with sorghum mtRNA.

Figure 11. Nucleotide sequence of the 5' flanking region of *atp9*. The *trn*fM sequence is shown in italics. The closed arrows represent the major termini mapped by primer extension analysis. The locations of the oligonucleotides are underlined. The open triangles represent the termini of protected fragments from *in vitro* capping in conjunction with ribonuclease protection experiments.

suggests that *atp9* was co-transcribed with *trn*fM and the post transcriptional processing event(s) generated heterogeneous transcript termini of *atp9*. In addition, *atp9* possessed an initiation region that initiated transcription at the positions -309, -305 and -302 nt (Figure 11).

Another primer, DP73 (Table 2), which is located 220 bp upstream of DP48 and 35 bp 3' to the 3' end of *trn*fM, was used to detect other primary transcripts. No extended fragments could be detected. The primary transcripts may be processed too rapidly to be detected by this assay.

Phosphorylation Reaction

The T4 polynucleotide kinase exchange reaction (Sambrook et al. 1989) has been reported to label only processed mtRNA, not initiated transcripts (Mulligan et al. 1988a, Auchincloss and Brown 1989), and this enzyme was used in this study. The termini other than at -309, -305, and -302 presumably resulted from processing events, and could be recognized by the T4 polynucleotide kinase. Hybridization of 5' phosphorylated mtRNA with a Southern blot containing subclone pRS66 DNA (Salazar 1991) is shown (*Pst*I-*Eco*RI, Figure 8E). Phosphorylated mtRNA hybridized with pRS66, suggesting that pRS66 includes 5' termini recognized by polynucleotide kinase.

Analysis of *trn*fM

To determine if there were other copies of *trn*fM located in the sorghum mitochondrial genome, a Southern blot containing restricted total mitochondrial DNA was probed with DNA fragments covering part (pRS72, Salazar 1991) or all of *trn*fM (Figure 12 B). MtDNA from line Tx398 was restricted with *Pst*I, *Sma*I and *Xba*I, respectively, and probed with an *Eco*RI-*Pst*I fragment of clone pRS72, which includes the 5' flanking

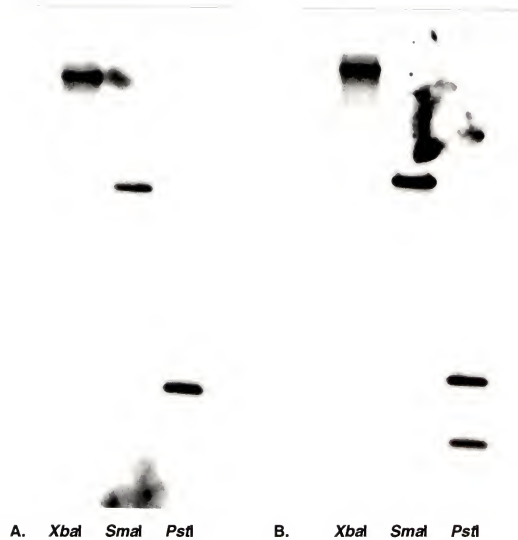


Figure 12. Hybridization of restricted mtDNA (*Xba*I, *Sma*I and *Pst*I) with pRS72 (A) and a fragment (B) covering the region from -195 to -496 5' to the *atp9* start codon (Figure 11). pRS72 covers the region from residue -344 to -799, relative to the *atp9* start codon (Figure 11). A single hybridization band was detected in A. In B, two hybridization bands were detected when the mtDNA was restricted with *Pst*I, since the probe includes a *Pst*I site (Figure 11).

and part of the coding sequence of *trn*fM, and probed with a fragment covering the coding sequence of *trn*fM. Hybridization showed only one band in each lane, suggesting that *trn*fM, located 5' to *atp9*, is the only copy of this gene in the genome of Tx398. Two hybridization signals were detected in Figure 12 B when mtDNA was restricted with *Pst*I, as the probe covers a *Pst*I site.

To further determine if the *trn*fM is transcribed, a DNA blot containing pRS72 DNA was hybridized to labeled small molecular weight RNAs (Figure 13). The small molecular weight RNAs were obtained by precipitation of LiCl soluble RNAs with ethanol, treated with DNAse and alkaline phosphatase, followed by polynucleotide kinase labeling. The *Sac*I-*Pst*I fragment of pRS72 (Figure 13) hybridized to the labeled probe, demonstrating that *trn*fM is transcribed. Hybridization of the *Eco*RI-*Sac*I fragment (Figure 13) may result from transcription of the *Eco*RI-*Sac*I region.

Transcription Initiation of *orf107*

Orf107 is a 321 bp chimeric open reading frame of the A3 male sterile cytoplasm. The chimeric configuration involves 310 bp 93% similar to sequences 5' and internal to *atp9*, including 31 amino terminal residues 84% similar to ATP9 (Tang et al. in preparation). Sequences 3' to the *atp9* similarity are of unknown origin. The carboxy terminal 54 residues are 52% identical to the carboxy terminus of *orf79*, an open reading frame associated with CMS in rice (Akagi et al. 1994). A total of four editing events were observed for *orf107* transcripts, altering four amino acids, including one within the *atp9* similarity region (Tang et al. in preparation).

Transcripts of *orf107* include species of about 1100, 870, 810 and 380 nt. The 1100, 870 and 810 nt transcripts are major transcripts in A3Tx398, while in line IS1112C,

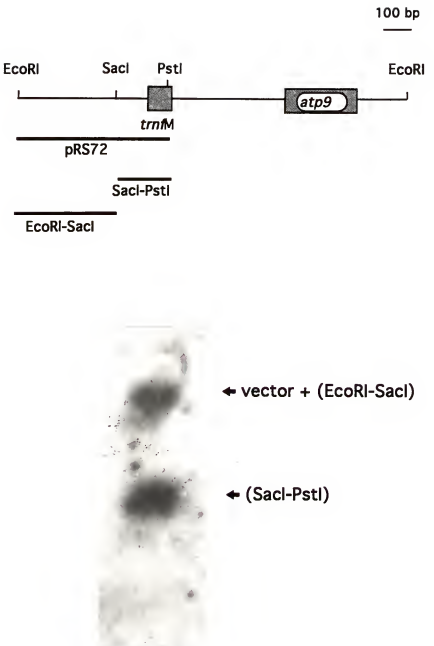


Figure 13. Hybridization of pRS72 DNA with [γ-32P]ATP labeled tRNA. *Sac*I and *Pst*I digested pRS72 DNA was electrophoresed in a 1% agarose gel, and blotted to nitrocellulose. The DNA blot was probed with labeled tRNA with polynucleotide kinase and [γ-32P]ATP.

a transcriptional processing event dramatically reduces the abundance of the 1100, 870 and 810 nt transcripts, and appearance of high levels of the 380 nt transcript.

Cloning Fragments 5' to *orf107* and Gene Specific Clones

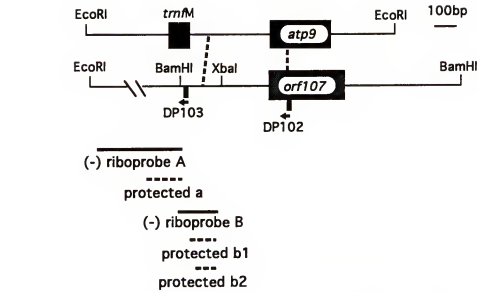
To clone a fragment 5' to clone pHC86 (Table 1), which contains *orf107* and the 5' flanking sequence up to -485 (Tang et al. in preparation), cosmid clone 8G8 (Salazar 1991) DNA was digested with *Bgl*II and probed with pHC86. A 2.9 kb *Bgl*II fragment was identified, cloned into *Bgl*II digested pUC19 vector (pBY124) and sequenced. This clone, pBY124, was used for subcloning a 1.8 kb *Eco*RI-*Bam*HI fragment into pUC19 (pBY128) as well as the Bluescript vector (pBY142) (Figure 14). Clone pBY142 is located 5' to clone pHC86.

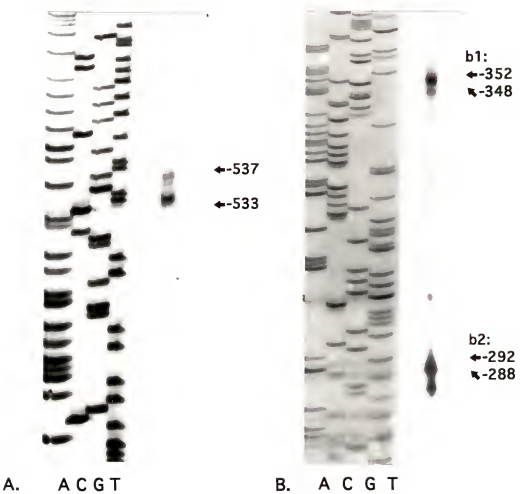
Another DNA segment, the 214 bp *Xba*I-*Bam*HI fragment of pHC103, was also cloned into the Bluescript vector (pBY145, Figure 14) and sequenced.

In vitro Capping and Ribonuclease Protection

Single-stranded antisense RNA was synthesized from *Eco*RI-digested pBY142 with T3 RNA polymerase, and from *Xba*I-digested pBY145 with T7 RNA polymerase, respectively. Hybridization of the capped RNA to the antisense RNA and ribonuclease digestion of the duplex resolved several protected fragments (Figure 14). In Figure 14A, antisense RNA used for protection was transcribed from pBY142 and two major protected fragments were detected, with termini corresponding to nucleotides of about -537 and -533 5' to the start codon. In Figure 14B, antisense RNA was transcribed from pBY145, and two protected regions were identified, with protected termini corresponding to -352, -348, -292 and -288 5' to the start codon.

Figure 14. Ribonuclease protection of the *orf107* transcript. Antisense RNAs (solid bar) synthesized from Bluescript clones pBY142 (riboprobe A) and pBY135 (riboprobe B) were hybridized with capped mtRNAs, and treated with ribonuclease. The protected regions (dashed line) were resolved with M13 sequence as a marker.


EcoRI
trnM
EcoRI
100bp
atp9
EcoRI
BamHI
XbaI
BamHI
orf107
DP103
DP102
(-) riboprobe A
protected a
(-) riboprobe B
protected b1
protected b2


-537
-533
b1:
-352
-348
b2:
-292
-288
A. A C G T
B. A C G T

Primer Extension

To identify precisely the 5' transcript termini of *orf107*, primer extension experiments were conducted. Oligonucleotide DP102 (Table 2), complementary to a sequence between +65 and +85, and is located 3' to the *atp9* similarity region, was prepared and used in primer extension with total mtRNA from line A3Tx398. Several major products were detected, with termini mapped to residues of -233, -237, -268, -292, -296, and -358 (Figure 15A), 5' to *orf107* (Figure 16). In addition, a longer fragment corresponding to the nucleotide of about -539 was also detected (not shown)

Another primer, DP103 (Table 2), complementary to a sequence of -364 to -383, relative to the start codon, and 460 bp 5' to DP102, was used to verify the site of the distal terminus obtained by primer DP102. The extended fragment was analyzed (Figure 15B). A major terminus was detected corresponding to residue of -539, upstream of the translational start codon (Figure 16).

Transcription Initiation of *urf209*

An open reading frame coding for 25 kD gene product, "*orf25*", has been identified in many plant specifies, including tobacco (Stamper et al. 1987), maize (Dewey et al. 1986, Stamper et al. 1987), wheat (Bonen et al. 1990), rice (Liu et al. 1992), *Marchantia* (Oda et al. 1992) and *Arabidopsis* (Brandt et al. 1992). The function of the gene product is unknown, and the open reading frame is variable in size in different species. In sorghum, *orf25* is designated as *urf209*, predicting 209 amino acid residues (Tang et al. submitted). A total of 10 C-to-U editing events were revealed, altering nine amino acids. Transcription analysis showed that there are two major transcripts of 1050 and 830 nt in IS1112C (Tang

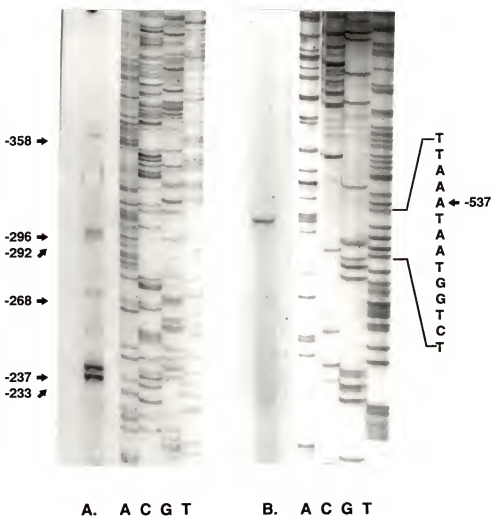


Figure 15. Primer extension analysis of *orf107*. A: Oligonucleotide DP102 was used in dideoxy sequencing of clone pBY142 and the primer extension reaction with sorghum mtRNA. B: Oligonucleotide DP103 was used in dideoxy sequencing of clone pBY142 and the primer extension reaction.

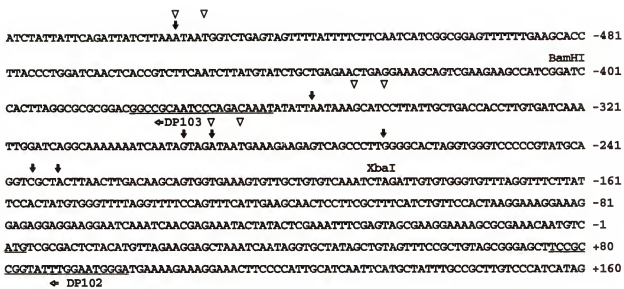


Figure 16. Nucleotide sequence of the 5' flanking region of *orf107*. The closed arrows represent the major termini mapped by primer extension analysis. The locations of the oligonucleotides are underlined. The open triangles represent the termini of protected fragments from *in vitro* capping in conjunction with ribonuclease protection experiments.

et al. submitted), and primer extension experiments identified the 5' termini of the *urf209* transcripts to residues of -345 and -134, 5' to the start codon.

Cloning Gene Specific Fragments

A gene-specific fragment covering the termini identified by primer extension, including 5' flanking and part of the coding sequences of *urf209*, was prepared from clone pHT98 digested with *Bgl*II and *Bam*HI, and cloned into *Bam*HI-restricted Bluescript vector. Sequencing revealed that clone pBY146 had the *Bgl*II end adjacent to the T3 promoter, and that the T3 RNA polymerase initiated antisense transcription of the cloned fragment.

In vitro Capping and Ribonuclease Protection

To identify if the termini determined by primer extension are primary ends, single stranded antisense RNA was synthesized from *Bam*HI-digested pBY146 with T3 RNA polymerase, and hybridized to the capped RNA from IS1112C. The protected fragment was analyzed (Figure 17). A major terminus was detected, corresponding to nucleotide of about -331 (Figure 18).

Transcription Initiation of *orf265*

Orf265 has been described as a chimeric gene specific to the A3 male-sterile sorghum cytoplasm, and resulted from a series of recombinational events (Tang et al., submitted). The 5' region of the open reading frame includes a 288 bp sequence 95% identical to the 3' terminal region of a repeat present 5' to maize T cytoplasm *atp6* and *T-urf13* (Dewey et al. 1986), including a common transcription initiation region for the two genes in T cytoplasm

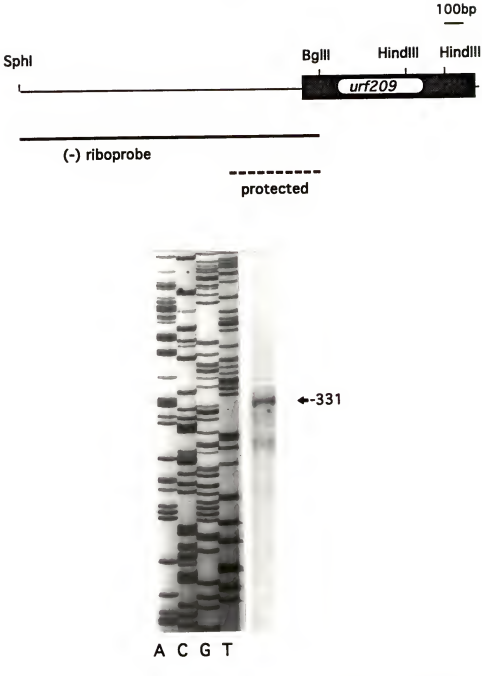


Figure 17. Ribonuclease protection of the *urf209* transcript. Antisense RNA (solid line) synthesized from Bluescript clone pBY146 was hybridized with capped mtRNA, and treated with ribonuclease. The protected region (dashed line) was analyzed with M13 sequence as a marker.

Figure 18. Nucleotide sequence of the 5' flanking region of *urf209*. The major terminus mapped by *in vitro* capping and ribonuclease protection is indicated by the open triangle. The closed arrow represents the terminus determined by primer extension (Tang et al. submitted).

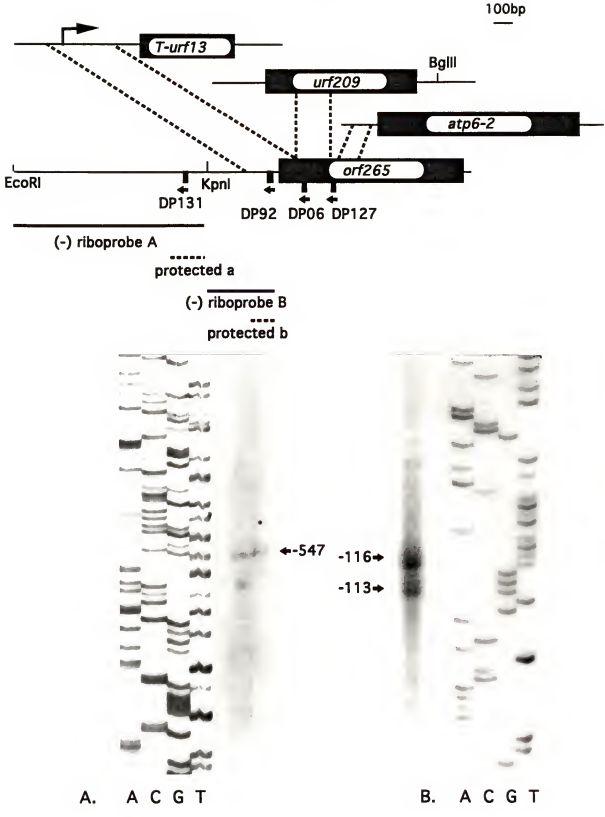
maize (Kennell and Pring, 1989). Within the maize-like sequence, an in-frame start codon is present followed by 53 bp of maize-like sequence and 189 bp of sorghum *urf209* . The insert corresponds to codons 51-114 of *urf209*. Sequences of 106 bp 3' to the *urf209* similarity region are identical to the sequence 5' to sorghum *atp6-2*, followed by sequences of unknown origin. An open reading frame of 795 bp was deduced in this configuration, encoding 265 amino acids. An open reading frame coding for 130 residues can be deduced if a second in-frame start codon was used. A schematic diagram of the *orf265* chimeric structure is shown in Figure 19. The chimeric *orf265* exhibits four base pair changes within the *urf209* similarity region, changing three predicted amino acids. Two transcripts of about 3500 and 3100 nt were detected by Northern analysis.

Cloning Gene Specific Fragments

To map the 5' transcript termini of *orf265*, the 5' flanking region of *orf265* was cloned. A 310 bp fragment, which is between -13 and -330, was amplified by PCR with primers DP92, 5'-GAGAGGCTACCAGGACACG-3' and the universal primer 1201, using pHT75 DNA (Tang et al. submitted) as a template. The amplified fragments were cloned into the Bluescript vector and sequencing revealed that clone pBY133 had the DP92 end adjacent to the T3 promoter, and that the T3 RNA polymerase initiated antisense transcription of the cloned fragment.

Another fragment which covered the region between -330 bp to about -1.1 kb from pHT100 (*Eco*RI-*Eco*RI) was cloned into the Bluescript vector, and sequencing revealed that clone pBY149 had the distal *Eco*RI end adjacent to the T3 promoter and that the T3 RNA polymerase initiated antisense transcription of the cloned fragment.

Figure 19. Ribonuclease protection of the *orf265* transcript. Antisense RNAs (shown in solid bar) synthesized from Bluescript clones pBY149 (riboprobe A) and pBY133 (riboprobe B) were hybridized with capped mtRNA, and treated with ribonuclease. The protected regions (dashed line) were analyzed with M13 sequence as a marker in (A) and with clone pBY148 sequence as a marker in (B).


100bp
T-urf13
urf209
BglII
atp6-2
orf265
EcoRI
KpnI
DP131
DP92
DP06
DP127
(-) riboprobe A
protected a
(-) riboprobe B
protected b
-547
-116
-113
A. A C G T
B. A C G T

Mapping Transcription Initiation Sites of *orf265*

Total mtRNA from line IS1112C was capped and hybridized with antisense RNAs synthesized from pBY149 and pBY133, respectively. The duplex was treated with ribonuclease, and the protected regions were analyzed. A terminus corresponding to a residue of about -547 (Figure 19A) was detected with the antisense RNA derived from clone pBY149. Using antisense RNA synthesized from clone pBY133, two major protected fragments were identified, with termini corresponding to residues of -116 and -113 (Figure 19B) 5' to the start codon.

Primer Extension

To precisely determine 5' termini of the *orf265* transcripts, primer extension analysis was conducted. Oligonucleotide DP127 (Table 2), which is complementary to a sequence of +296 to +278 5' to the start codon, and located 3' to the *urf209* similarity region, was prepared for primer extension analysis. A major terminus corresponding to residue -116 was identified (Figure 20A). In addition, a longer product was also detected, with terminus corresponding to the nucleotide of -554 (not shown).

Primer DP06 (Table 2), which is complementary to a sequence of +131 to +112 and located within the *urf209* similarity region (Figure 19), was also used in primer extension analysis and gave a result consistent with DP127 (Fig. 20B). Since primer DP06 is located within *urf209* similarity region, it may detect the termini of *orf265* transcripts as well as *urf209* transcripts. Another oligonucleotide DP131 (Table 2), complementary to a sequence of -31 to -51, relative to the start codon, and is located upstream of the *urf209* similarity region (Figure 19), was used to map the longer transcript terminus. A major terminus was detected corresponding to the nucleotide of -554 upstream of the *orf265* start codon (Figure 20C, Figure 21).

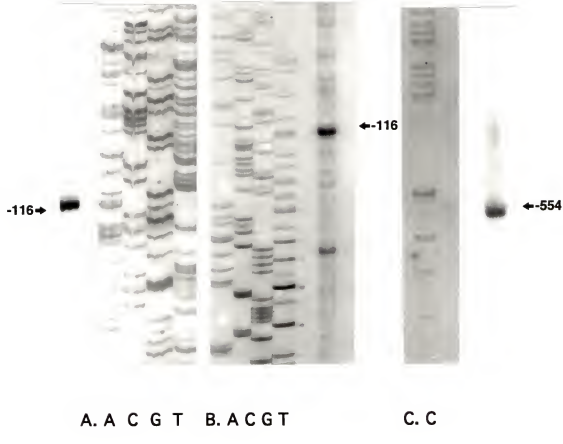


Figure 20. Primer extension analysis of *orf265*. A: Oligonucleotide DP127 was used for the primer extension reaction with clone pBY146 sequence as a marker. B: Oligonucleotide DP06 was used in primer extension with M13 sequence as a marker. C: Oligonucleotide DP131 was in primer extension with the C track of M13 sequence as a marker.

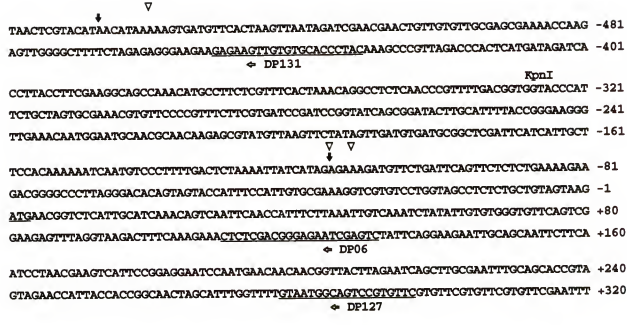


Figure 21. Nucleotide sequence of *orf265* and the 5' flanking region. The closed arrows represent the major termini mapped by primer extension analysis. The locations of the oligonucleotides are underlined. The open triangles represent the termini of protected fragments from *in vitro* capping in conjunction with ribonuclease protection experiment.

In vitro Transcription of *atp6-2* with Maize Extract

Plasmid pBY148 insert (*Eco*RI-*Bgl*II) covering a region from -38 to -338 (Figure 7) 5' to the *atp6-2* start codon was programmed with maize transcriptional extract. A transcript of 160 nt was detected (Figure 22), corresponding to the initiation at the -188 nucleotide (Figure 6) and the termination at the end of the template (*Bgl*II site), demonstrating that the sequence surrounding -188 nt was specifically recognized by RNA polymerase. Thus the DNA element around -188 is referred to as a promoter.

Figure 22. *In vitro* run-off transcript analysis. Clone pBY148 was digested with *Eco*RI and *Bgl*II to release the insert, and the insert was used as a template in an *in vitro* transcription reaction with maize extract. Products were analyzed with M13 sequence as a marker.

335 nt
160 nt
A C G T

DISCUSSION

Transcript patterns of plant mitochondrial genes are usually complex. Frequent recombinations and rearrangements may shuffle a promoter region and result in chimeric sequences or loss of promoter function (Gray et al. 1992). Initiation of transcription may occur at one or several promoters for one gene (Gray et al. 1992). Furthermore, extensive transcriptional processing may generate heterogeneous 5' transcript termini that can superimpose a further complexity to the transcript patterns of plant mitochondrial genes. Alignments of sequences surrounding transcription initiation sites of plant mtDNA revealed a limited conserved YRTA motif (Rapp and Stern 1992). The YRTA block has been tested in an *in vitro* transcription system and shown to be an essential element for the promoter function of maize *atp1* (Rapp and Stern 1992, Rapp et al. 1993). However, DNA sequences around transcription initiation sites for some plant mitochondrial transcripts show little similarity to the YRTA element (Binder et al. 1994, Lizama et al. 1994, Mulligan et al. 1991). Taken together, these characteristics may indicate that the plant mitochondrial transcription complex may not have a stringent sequence recognition requirement.

To provide more information about plant mitochondrial transcription initiation, we investigated 5' transcript termini of *Sorghum bicolor*. The objectives of the study include characterization of transcription initiation of selected sorghum mitochondrial genes and open

reading frames as a preface to elucidate mitochondrial gene regulation at the transcriptional level. The availability of basic transcriptional data and cloned 5' sequences allowed the examination of nine transcription initiation regions for four genes and two open reading frames, which should give a good representation of transcriptional initiation in sorghum mtDNA.

The nine initiation regions identified in this study by ribonuclease protection are, generally, in good agreement with the data by primer extension experiments. This concordance are interpreted as evidence that the number and the position of the transcription initiation sites identified through ribonuclease protection experiments are consistent with the results obtained by primer extension analyses. Assignments of the 5' termini of the protected RNA fragments to the corresponding DNA sequence were based on sizing of the protected fragments. The sizing of protected RNA was approximated by electrophoresis using M13 or mtDNA genomic clones as markers, taking the RNA and DNA mass differences into account. Several protected fragments were sized somewhat shorter than predicted by primer extension data. Instability at the 3' terminus of the DNA-RNA heteroduplex due to specific nucleotide sequences, which could allow ribonuclease hydrolysis, might explain these discrepancies. Considering the complexities of sizing protected fragments, results from primer extension experiments were thus used in promoter sequence alignments in this investigation.

On the other hand, experiment artifacts can be obtained by primer extension through unstable template-enzyme complexes influenced by secondary structure or sequences that inhibit stability, resulting in artificial 5' transcript termini. Moreover, sequences 5' to some plant mitochondrial genes may share sequence similarities (Pring et al. 1992), and thus an oligonucleotide may anneal to sequences 5' to other genes. Taking

these possibilities into considerations, most primer extension experiments in this investigation were performed with more than one oligonucleotide to confirm each result.

The Transcription Initiation Regions of *atp6-1* and *atp6-2*

Consistent results of mapping 5' transcript termini of sorghum *atp6-1* were obtained by ribonuclease protection and primer extension experiments, as exemplified by the number and the positions of the termini identified. The two 5' transcript termini of *atp6-1* were assigned to the corresponding DNA sequence of -441 and -437 by *in vitro* capping and ribonuclease protection, and to residues of -442 and -439 by primer extension. Based on the concordance of the results by two different approaches of protection and primer extension, transcription of *atp6-1* was assumed to initiate at two initiation sites of -442 and -439. The tentative 3' transcript terminus of *atp6-1* is at +1103 (D. R. Pring and F. Kempken, unpublished data), and thus initiation of transcription at -442 and -439 generates transcript of about 1540 nt, which is concordant with the result obtained from Northern analysis (Mullen et al. 1992). Analysis of sequences surrounding the transcription initiation sites at -442 and -439 revealed that the *atp6-1* promoter lacks a YRTA core element. Instead, a CTTA element was found, with initiation at +5 and +8 relative to the CTTA.

Two protected termini were detected for *atp6-2* by ribonuclease protection experiments. Primer extension analyses assigned the 5' terminus of the *atp6-2* transcript to -188, concordant with the result of the long protected fragment. Since primer extension detected only one terminus, in contrast of two termini identified by protection, the 5' transcript terminus of *atp6-2* was tentatively assigned to the terminus of the long protected fragment, which is the identical terminus detected by primer extension. The short terminus

might result from ribonuclease digestion extending into 3' terminus of the protected fragment.

DNA sequence comparisons showed that the transcription initiation region of sorghum *atp6-2* is highly similar to a maize *atp6* promoter (Rapp and Stern 1992), including CGTA element. A sequence beginning 36 bp 5' to *atp6-2* is identical to chloroplast *rrn23* (Pring et al. 1993). The mitochondrial transcriptional apparatus apparently recognized the sequence 3' to the chloroplast *rrn23* and can initiate transcription of *atp6-2*. This suggests that the 36 bp immediately upstream of the initiation site may be sufficient for the RNA polymerase recognition. It may also be argued that the chloroplast sequence retained sequence element(s) that could be recognized by mitochondrial transcription complex.

The Transcription Initiation Region of *atp9*

Transcription of *atp9* is characterized by a single transcript of 650 nt (Salazar, 1991). Since a minimum transcript length of 636 nt is required to cover both *trn*fM and *atp9* to the edited stop codon, the 650 nt transcript is assumed to be the mature transcript of *atp9* (Salazar, 1991). Given that all plant mitochondrial *trn*fM identified are mitochondrion-encoded (Marechal-Drouard et al. 1993), and there is only one copy of *trn*fM in the Tx398 genome, this copy of *trn*fM is expected to be a functional gene.

The assigned primer extension termini and the termini suitable for guanylyltransferase of *atp9* represent a novel observation. In other similar systems with *trn*fM present 5' to a protein coding gene, excision of *trn*fM has been postulated to be a precise endonucleolytic cleavage at the 3' terminus of *trn*fM, generating the mature 3' end of *trn*fM and a 5' terminus of the 3' transcript (Singh and Brown 1991). The multiple

termini of sorghum *atp9* primer extension may represent multiple endonucleolytic cleavages producing multiple 5' termini of the *atp9* transcripts. Alternatively, multiple transcript termini may result from primer extension artifacts, as described previously. Interactions of the *atp9* transcripts and the AMV reverse transcriptase may be influenced by secondary structures of the *atp9* transcripts and lead to unstable template-enzyme complex. This unstable complex may result in the dissociation of the reverse transcriptase from the template-enzyme complex.

Surprisingly, some of the *atp9* transcript termini were substrates for guanylyltransferase, with a 5' sequence 5'-*T*AATCACGTCAGATACTTGTACGAGTCCTGATCA-3', bearing a YRTA core motif (**TGTA**, boldface. The 3' end of *trn*fM is shown as the italicized T). The initiation sites (single-underlined nucleotides, plus the C residue adjacent to the TGTA core) were assigned to nucleotides of -309, -305 and -302. Given the complex pattern of the *atp9* primer extension, only approximate assignment, rather than unambiguous positioning, of the 5' initiated termini to the corresponding primer extension termini was possible. Since all transcription initiation sites identified for plant mtDNA were positioned 3' to the core YRTA block (Hanic-Joyce and Gray 1991, Rapp and Stern 1992), the transcription initiation sites of *atp9* were thus assigned to residues of -302, -296 or -295, and -291 (double-underlined residues of above sequence) with the initiation sites at +1, +7 or +8, and +12, agreeing with primer extension data.

Assuming the termini suitable for guanylyltransferase were primary transcript ends, the novel observation of initiation of transcription at a region 3' to *trn*fM suggests that sequences 5' to the initiation site, as few as about 21 bp, may be sufficient for the recognition of RNA polymerase, with certain 3' sequence elements. Alternatively, apart from the 21 bp DNA segment, some 5' sequences corresponding to part of *trn*fM coding

region may be involved as a promoter of *atp9* and play a role in the initiation of transcription.

Transcript termini of sorghum mitochondrial *atp9* may therefore be characterized by two origins: processing of the precursor at the a region 3' to *trn*fM, resulting in 5' processed transcript termini; and uniquely, initiation of transcription 3' to the *trn*fM, generating 5' primary *atp9* transcripts.

The observation of unusual processing generating substrates suitable for guanylyltransferase revealed the possibility that capped termini may not be necessarily associated with initiation termini in some circumstances (L'Abbe et al. 1990). As a result, the putative transcription initiation region of *atp9* identified in this investigation need to be examined in an *in vitro* transcription system.

The Transcription Initiation Regions of *orf107*

Transcripts of *orf107* in A3Tx398 include species of about 1100, 870, and 810 nt (Tang et al., in preparation), and initiation of transcription of *orf107* was determined to occur at residues of -537, -533, -352, -348, -292 and -288, 5' to the start codon. Compared with the data obtained by primer extension, which includes termini at -537, -358, -296, -292, -268, -237 and -233, transcript termini at -268, -237 and -233 were assigned to the processed transcripts while other termini corresponded to the initiated ends. The distal terminus, at -537 by primer extension, is correlated with the two protected fragments at -537 and -533. The middle terminus, at -358 determined by primer extension corresponds to protection fragments at -352 and -348. The two proximal termini at -296 and -292 by primer extension are in agreement with two termini of -292 and -288 by protection, and thus initiation at sites -296 and -292 is proposed. Taken together, three

initiation regions-- the distal, middle and proximal regions--were identified for *orf107* transcripts, with single initiation sites at -537 and -358 for the distal and middle region, respectively, and two initiation sites at -296 and -292 for the proximal initiation region.

Alignments of the sequences surrounding the initiation sites indicate that the middle region is in good agreement with the YRTA (TATA) motif with initiation at +3 relative to the TATA element. The other two initiation regions, however, do not fit the YRTA consensus. Instead, they contain a CTTA or AATA with initiation at +1 and +2 respectively.

The Transcription Initiation Region of *urf209*

A single protected fragment was detected by ribonuclease protection corresponding to a single terminus identified by primer extension at -345 (Tang et al. submitted). There is a 14 bp difference between the data of -331 (Figure 17) of protection and -345 by primer extension. This difference was the largest discrepancy among examples described, and may have resulted from possible ribonuclease digestion extending into the 3' terminus of the protected fragment. The data do suggest that a single terminus is associated with a single initiation site, which was assigned to the terminus of 1050 nt transcript of *urf209*. The primer extension data was thus used for sequences comparisons. Alignments of sequences surrounding the initiation site, assigned to -345, revealed that the initiation region lacked a YRTA motif, and instead, a AATA element was present, as shown for the proximal initiation region of *orf107*.

The Transcription Initiation Regions of *orf265*

Two transcription initiation regions were identified for *orf265*, which are about 400 bp apart, consistent with the analysis of steady state transcripts of 3500 and 3100 nt (Tang et al. submitted). The transcript terminus at -116 (underlined nucleotide), 5'-ACTCT*AAA*ATT*AT*CATAGAGAAAGA-3', is within a maize-like sequence, and resides in a promoter region (the proximal promoter) for *T-urf13* in maize (Kennell and Pring 1989). Comparing the sequences surrounding the initiation site of *orf265* in sorghum and T-*urf13* in maize, some base substitutions were found (italicized nucleotides). The substitutions at -2, and -6 to -8, relative to the YRTA core, fall into the N regions of a consensus, A__NNYNRYRTAT, derived for maize *atp1* and other maize promoters (Rapp et al. 1993). Consequently, the substituted DNA segment still retained a promoter function.

A sequence of 288 bp 5' to sorghum *orf265* is 95% identical to a maize repeated sequence present 5' to maize *T-urf13* and *atp6* (Tang et al. submitted). Sequences 5' to the recombination break point of *orf265*, 5'-GAAACAATGGAATGCAACGCAACAAGAG***CGTATGTTAAGTTCTATAGT-3'*** (the sequences shared by maize and sorghum are in italics) share some similarities to the maize sequence, 5'-ATTGATGGGTAATAGTTTTGATAGAAAACT***CGTATGTTAAGTTCTCTAGT***-3' This maize sequence includes the middle promoter of *T-urf13* and *atp6* in maize (deduced from Rapp and Stern 1992; and Dewey et al. 1986). The corresponding *orf25* sequence is not active in transcription initiation, although the sorghum DNA region includes CRTA (boldface) from the maize-like sequence and repeats of GAA or CAA (underlined) present 5' to the break point. This suggests other 5' domains in addition to the core CRTA may be

required for the promoter function. Another possibility that cannot be ruled out is that an adenosine substitution of cytosine at +12 (not shown) relative to the YRTA core motif might abolish promoter function, since sequences 3' to the 11 nt sequence have not been tested for effects on promoter function (Rapp and Stern 1992).

A distal initiation region was also identified for orf265, with a sequence 5'-GAATTAACTCGTACATAA-3', surrounding the initiation site -554. This sequence block includes a perfect CRTA and upstream GAA repeat, with initiation at +4, concordant with proposed promoter motifs for other species (see below).

In vitro Transcription of *atp6-2*

Among the nine regions of transcription initiation identified in this study, a sequence, 5'-GAAAACTCGTATAATAG-3', covering the initiation site (underlined) of *atp6-2* is highly similar to the maize *atp6* promoter, and thus a DNA fragment bearing the *atp6-2* transcription initiation region was tested in a heterologous *in vitro* transcription system with maize extract. Partially purified RNA polymerase from maize mitochondria recognized the promoter of sorghum *atp6-2*, and initiated at a position consistent with the -188 site assigned from primer extension, and the longest protected fragment from *in vitro* capping and protection experiments (Figure 22). Heterologous *in vitro* transcription studies have been conducted in animal and fungi systems (Clayton 1991, Costanzo and Fox 1990), and some compatible or incompatible interactions have been revealed (Clayton 1991, Costanzo and Fox 1990). No such studies have been reported in plant mitochondrial systems to date. That the maize RNA polymerase recognizes a sorghum promoter suggests the transcription machinery of the two genera may be related.

Consensus Motif of Sorghum mtDNA Promoters

In an effort to ascertain sequence similarities between different initiation regions of the sorghum mitochondrial genes and open reading frames, sequences surrounding the initiation sites were compared and aligned (Figure 23). A consensus block of YRTA was apparent for five of the nine representatives, indicating some functions of the conserved motif in promoting transcription initiation of sorghum mtDNA, as described in other systems (Mulligan et al. 1991, Rapp and Stern 1992, Rapp et al. 1993). All mitochondrial promoters analyzed to date, including those of animals, yeast and plants, with one exception of the HSP in mouse (Chang et al. 1986b), are characterized by a DNA segment that covers the transcription initiation site (Biswas et al. 1990, Clayton 1991, Rapp and Stern 1992). Furthermore, the corresponding elements identified in maize by mapping of 5' transcript termini *in vivo* (Mulligan et al. 1991) have been demonstrated to be indeed critical for promoting transcription initiation *in vitro* (Rapp and Stern 1992, Rapp et al. 1993), as well as for protein binding (Newton et al. 1995). Consequently, the sequence elements described here are reasonably referred to as part of the putative promoters of sorghum mtDNA, although definitive characterization awaits the examination of these DNA elements in an *in vitro* transcription system.

The sorghum mitochondrial putative promoter sequences were aligned relative to the YRTA core of Rapp and Stern (1992) and Rapp et al. (1993). Five of the nine promoters bear the YRTA element, including promoters of obligate genes *atp6-2* and *atp9*. The four promoters that do not fit YRTA include the promoter of *urf209* and the proximal promoter of *orf107*, which exhibit an AATA (ARTA) motif. The promoter of *atp6-1* and the distal promoter of *orf107*, which display a CTTA motif, also deviate from the consensus sequence. The initiation sites range from +1 to +5 for most of the examples, with the

		-29 -20 -10 -1+1 5 10 14	ST
Zmatp1		TTGTTATTATTAAGCCGAAAAGTAA CGTA TTAAAAACAAAGT	+1
*ZMatp6/T-urf13*A		CTGTAAGCATCTTCCCATGAATTGA CGTA TATAAGCGTAGTTA	+4
*ZMatp6/T-urf13*B		TGGGTAATAGTTTTGATAGAAAACT CGTA TGTTAAGTTCTCT	+3
*ZMatp6/T-urf13*C		GGTTCCCTTTTTACTCTCCCATTTT CATA GAGAAAGATGTTCT	+5
Sbatp6-2:		TATCATGTTTGCTCCTCAGAAAACT CGTA TAATAGTCTCATTG	+5
Sborf107:	a)	TTTTAATCTATTATTCACAGATTAT CTTA AATAATGGTCTGAG	+2
	b)	GCGGACGGCCGCAATCCCAGACAAA TATA TTAATAAAGCATCC	+3
	c)	TCAAATTGGATCAGGCAAAAAAATC AATA GTAGATAATGAAAG	+1, +5
Sburf209:		CTAGATGGTATTTGGAGAATATCCT AATA AGTTTAGCTTATTT	+2
Sborf265:	a)	GAAAATAGGCACGCACGAATTAACT CGTA CATAACATAAAAAG	+4
	b)	ATGTCCCTTTTGACTCTAAAATTAT CATA GAGAAAGATGTTCT	+2
Sbatp9:		GCCCCGCCTAATCACGTCAGATACT TGTA CGAGTCCTGATCAC	1, 7/8, 12
Sbatp6-1:		TCGACTCTACGTCTACGTGAATAAA CTTA CAATGGAGATCTAC	+5, +8

```
maize consensus:              A/TYRTAG/TAA/TAAA
wheat consensus:            RAaaNNGCRTAtARtRagt
consensus for dicots:       RAaAtNtCRTAaGaGaGAagNaaR
yeast consensus:              ATATAAGTA
```

Figure 23. Sequences surrounding the transcription initiation sites of selected sorghum mtDNAs and the promoter consensus of mtDNAs for other species. Transcription initiation sites are underlined. The conserved YRTA sequence is underlined. ST stands for the site(s) of transcription, relative to the YRTA. R indicates purine, and Y stands for pyrimidine. Zmatp1 represents the promoter for maize atp1 (Rapp and Stern 1992). Zmatp6/T-urf13A, Zmatp6/T-urf13B, and Zmatp6/T-urf13C represents the distal, middle and proximal promoter of *atp6* and *T-urf13* in maize, respectively (Kennell et al. 1989, Rapp and Stern 1992). Sborf107a, Sborf107b and Sborf107c indicates the distal, middle and proximal initiation region of sorghum *orf107*, respectively. Sborf265a and Sborf265b represents the distal and the proximal initiation region of sorghum *orf265*, respectively.

second *atp6-1* site at +8, and the second and third sites of *atp9* at +7 and +8. Assignment of the core of the proximal promoter of *orf107* could be moved three bases 3' to the AATA (ARTA), yielding an AGTA (ARTA), but one of the two initiation sites would be assigned to the G residue of internal AGTA block, which is inconsistent with other proposed positioning relative to the core (Rapp and Stern 1992, Rapp et al. 1993).

The core domain of YRTA has been found to be necessary, but insufficient to promote transcription of cloned fragment by *in vitro* analysis of the maize *atp1* promoter (Rapp and Stern 1992). Other sequence element(s), such as the repeated elements G(A)3-4 or C(A)3-4, present just upstream of the transcription initiation site, as found in several plant systems, may be required for specific transcription initiation. Although the guanine preceding $(A)_n$ has been shown to not be essential *in vitro* for maize atp1 promoter function as demonstrated by mutating the GAAAA element to TAAAA (Rapp and Stern 1992), the GAAAA could play a role *in vivo* for promoting transcription initiation or the critical elements may reside within the AAAA domain. This upstream domain was positioned approximately 10 bp 5' to the core motif, corresponding to one helical turn (Rapp et al. 1993). The nine proposed promoters described here each exhibits elements of this domain, ranging from (G/T)AAAA to GA, and the domain is centered around -10, relative to the core YRTA.

The two-domain-model--one core motif, one upstream element--has been proposed for the maize *atp1* promoter (Rapp et al. 1993). Based on the data of the nine putative promoters identified, sorghum mtDNA promoters are expected to bear similar two domains. Inspection of sequences between the two domains of the nine sorghum promoters as well as promoters of maize and wheat revealed that sequences between the two domains lack nucleotide similarity, but rather a pyrimidine region with one to several nucleotides of C or/and T is always present. Conversely, a DNA fragment with two perfect domains and a

pyrimidine region in between may not be a promoter, since other factors may be involved in initiation transcription, in addition to the putative necessary elements. This circumstance is represented at the recombinational break point of the 288 bp sorghum-maize region in sorghum *orf265*. Sequences surrounding the break point resemble the middle promoter of *T-urf13* in maize (deduced from Rapp and Stern 1992 and Dewey et al. 1986), including a CGTA core motif residing in the maize-like sequence and a upstream domain of CAA and GAA from sorghum sequence region, but no pyrimidine in between. This chimeric region surrounding the break point has been shown to not be active as a promoter for *orf265*. Nevertheless, a possible role of this pyrimidine region, based on inspection, can only be concluded by testing in an *in vitro* transcription system.

In contrast to the systems in *Xenopus laevis* (Bogenhagen and Romanelli 1988), *Neurospora crassa* (Kubelik et al. 1990), yeast (Marczynski et al. 1989) and T7 bacteriophage (Klement et al. 1990), in which transcription initiates at a single position, multiple initiation sites at one promoter may exist for plant mitochondrial DNA (Rapp and Stern 1992). Closely-spaced multiple initiation sites were also found for the sorghum promoters. For instance, the promoter of *atp6-1* and the proximal promoter of *orf107*, and probably the *atp9* promoter, bear two initiation sites.

Among the nine putative promoters identified through this investigation, four promoters deviate from the consensus core motif, as observed in other species (Binder and Brennicke 1994, Lizama et al. 1994, Mulligan et al. 1991, Newton et al. 1995). These sequence-diverged promoters have been proposed to be weak promoters based on the run-on experiments (Mulligan et al. 1991). Data from the present study do not allow a conclusion regarding promoter strength between the four diverged promoters versus the five conserved promoters, without performing run-on experiments.

Comparative analyses of the middle and the proximal promoters of *orf107* with one protected fragment showed that the abundance of the protected fragment for the middle promoter was less abundant than that of the proximal promoter (Figure 14), although the core region of the middle promoter bears a YRTA (TATA) while the proximal promoter does not. Transcripts corresponding to these promoters do not illustrate convincing differences in abundance (Tang et al. in preparation). This suggests that there may not be necessary correlations between the abundance of the capped transcript and the abundance of steady state transcripts.

The lack of similarity of some plant mitochondrial promoters might be explained if plant mitochondrial transcription complex occasionally employs a scanning mechanism for transcription initiation so that the polymerase initiates at sites that possess only limited identity to the polymerase or factor binding motif (Brown et al. 1991). Alternatively, transcription initiation may involve more than one type of promoter. For instance, not all cappable termini that have the YRTA motif are suitable sites for *in vitro* RNA polymerase initiation, as demonstrated by *in vitro* analysis of the *rrn18* promoter (Rapp and Stern 1992). The maize *rrn18* promoter fits the proposed consensus perfectly, but was not recognized by the maize extract that correctly promotes transcription of protein-coding genes (Rapp and Stern 1992). Thus, a complex transcription machinery may exist for plant mtDNA.

Overall, this investigation has documented nine 5' primary transcript termini of sorghum mtDNA by demonstrating that each is a suitable substrate for guanylyltransferase. The defined sequence similarities located 5' to the transcriptional initiation sites, revealed by sequence comparisons, allow the identification of nine putative sorghum mitochondrial promoters. These nine promoters exhibit a consensus of two domains as proposed for the maize *atp1* promoter: the core motif and the -10 bp upstream domain. Five of the nine

promoters have a conserved core motif of YRTA as identified in other plant species (Mulligan et al., 1988a, 1988b, Covello et al., 1991, Hanic-Joyce and Gray, 1991, Brown et al., 1991, Rapp and Stern, 1992, Binder & Brennicke, 1993), while the other four bear only limited conserved motif of ARTA or YTTA. These observations suggest that plant mitochondrial promoters may indeed be characterized by a certain degree of flexibility of promoter sequences, concordant with complex transcription machinery.

Substantiation of the capability of the putative promoters to function in transcription initiation would be obtained through introduction of the elements in a transient gene expression system, *in vitro* transcription assays, or ultimately, into a transformation system. The development of *in vitro* RNA polymerase assays in wheat (Hanic-Joyce and Gray 1991) and maize (Rapp and Stern 1992) allows an assessment of the putative sorghum mitochondrial promoter. The sorghum *atp6-2* template was a suitable template for the partially purified maize RNA polymerase. The initiation at the site predicted by primer extension and *in vitro* capping suggests that the identified site is indeed a promoter.

REFERENCES

Akagi, H., Sakamoto, M., Shinjyo, C., Shimada, H., and Fujimura, T. 1994. A unique sequence located downstream from the rice mitochondrial *atp6* may cause male sterility. Curr. Genet. 25:52-58.

Altman, S. 1981. Transfer RNA processing enzymes. Cell 23:3-4.

Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R., and Young, I. G. 1981. Sequence and organization of the human mitochondrial genome. Nature 290:457-465.

Andre, C, Levy, A., and Walbot, V. 1992. Small repeated sequences and the structure of plant mitochondrial genomes. Trends in Genetics 8:128-132.

Attardi, G., and Schatz, G. 1988. Biogenesis of mitochondria. Ann. Rev. Cell Biol. 4:289-333.

Auchincloss, A. H., and Brown, G. 1989. Soybean mitochondrial transcripts capped *in vitro* with guanylyltransferase. Biochem. Cell. Biol. 67:315-319.

Bailey-Serres, J., Dixon, L. K., Liddell, A. D., and Leaver, C. J. 1986a. Nuclear-mitochondrial interactions in cytoplasmic male-sterile sorghum. Theor. Appl. Genet. 73:252-260.

Bailey-Serres, J., Hanson, D. K., Fox, T. D., and Leaver, C. J. 1986b. Mitochondrial genome rearrangement leads to extension and relocation of the cytochrome c oxidase subunit I gene in sorghum. Cell 47:567-576.

Bendich, A. 1993. Reaching for the ring: the study of mitochondrial genome structure. Curr. Genet. 24:279-290.

Binder, S., and Brennicke, A. 1993a. A tRNA gene transcription initiation site is similar to mRNA and rRNA promoters in plant mitochondria. Nucl. Acids Res. 21:5012-5019.

Binder, S., and Brennicke, A. 1993b. Transcription intiation sites in mitochondrai of *Oenothera berteriana*. J. Biol. Chem. 268:7849-7855.

Binder, S., Thalheim, C., and Brennicke, A. 1994. Transcription of potato mitochondrial 26S rRNA is initiated at its mature 5' end. Curr. Genet. 26:519-523.

Biswas, T. K., and Getz, G. S. 1990. Regulation of transcription initiation in yeast mitochondria. J. Biol. Chem. 265:19053-19059.

Boeshore, M. L., Hanson, M. R., and Ishar, S. 1985. A variant mitochondrial DNA arrangement specific to Petunia stable variant somatic hybrids. Plant Mol. Biol. 4:125-132.

Bogenhagen, D., and Romaneli, M. F. 1988. Template sequences required for transcription of *Xenopus laevis* mitochondrial DNA from two bidirectional promoters. Mol. Cell. Biol. 8:2917-2924.

Bonen, L., Bird, S. and Belanger, L. 1990. Characterization of the wheat mitochondrial *orf25* gene. Plant Mol. Biol. 15: 793-795.

Bonen, L., and Brown, G. 1993. Genetic plasticity and its consequences: perspectives on gene organization and expression in plant mitochondria. Can. J. Bot. 71:645-660.

Bonen, L. Williams, K., Bird, S., and Wood, C. 1994. The NADH dehydrogenase subunit 7 gene is interrupted by four group II introns in wheat mitochondrial genome. Mol. Gen. Genet. 244:81-90.

Bonnard, G., Gualberto, J. M., Lamattina, L.,and Grienenberger, J. M. 1992. RNA editing in plant mitochondria. Critical Reviews in Plant Sciences 10:503-524

Bonnard, G. and Grienenberger, J. M. 1995. A gene proposed to encode a transmembrane domain of an ABC transporter is expressed in wheat mitochondria. Mol. Gen. Genet. 246:91-99.

Bonhomme, S., Budar, F., Lancelin, D., Small, I., Defrance, M. C., and Pelletier, G. 1992. Sequence and transcript analysis of the *Nco2.5* Ogura-specific fragment correlated with cytoplasmic male-sterility in *Brassica* cybrids. Mol. Gen. Genet. 235:340-348.

Brandt, P., Sunkel, S., Unseld, M., Brennicke, A. and Knoop, V. 1993. The *nad4L* gene is encoded between exon c of *nad5* and *orf25* in the *Arabidopsis* mitochondrial genome. Mol. Gen. Genet. 236:33-38.

Braun, C. J., Brown, G. G., and Levings, C. S., III. 1992. Cytoplasmic male sterility. In: Control of plant gene expression, p219-245, D. P. S. Verma, ed. CRC Press, Boca Raton.

Brettell, R. I. S., Thomas, E., and Ingram, D. S. 1980. Reversion of Texas male-sterile cytoplasm maize in culture to give fertile, T-toxin resistant plants. Theor. Appl. Genet. 58:55-58.

Brown, G. G., Auchincloss, A. H., Covelo, P. S., Gray, M. W., Menassa, R., and Singh, M. 1991. Characterization of transcription initiation sites on the soybean

mitochondrial genome allows identification of a transcription-associated sequence motif. Mol. Gen. Genet. 228:345-355.

Chang, D. P., and Clayton, D. A. 1984. Precise identification of individual promoters for transcription of each strand of human mitochondrial DNA. Cell 36:635-643.

Chang, D. D., and Clayton, D. A. 1986a. Precise assignment of the light strand promoter of mouse mitochondrial DNA: a functional promoter consists of multiple upstream domains. Mol. Cell. Biol. 6:3253-3261.

Chang, D. D., and Clayton, D. A. 1986b. Precise assignment of the heavy strand promoter of mouse mitochondrial DNA: cognate start sites are not required for transcriptional initiation. Mol. Cell. Biol. 6:3262-3267.

Chase, C. D., and Pring, D. R. 1986. Properties of the linear N1 and N2 plasmid-like DNAs from mitochondria of cytoplasmic male-sterile *Sorghum bicolor*. Plant Mol. Biol. 6:53-64.

Chaumont, F. Bernier, B., Buxant, R., Williams, M., Levings III, C., and Boutry, M. 1995. Targeting the maize *T-urf13* product into tobacco mitochondria confers methomyl sensitivity to mitochondrial respiration. Proc. Natl. Acad. Sci. (USA) 92:1167-1171.

Chen, Z., Muthukrishnan, S., Liang, G. H., Schertz, K. F., and Hart, G. E. 1993. A chloroplast DNA deletion located in RNA polymerase gene *rpoC2* in CMS lines of sorghum. Mol. Gen. Genet. 236:251-259.

Christianson, T., and Rabinowitz, M. 1983. Identification of multiple transcription initiation sites on yeast mitochondrial genome by *in vitro* capping with guanylyltransferase. J. Biol. Chem. 258:14025-14033.

Clayton, D. A. 1984. Transcription of mammalian mitochondrial genome. Annu. Rev. Biochem. 53:573-94

Clayton, D. A. 1991. Replication and transcription of vertebrate mitochondrial DNA. Annu. Rev. Cell Biol. 7:453-478.

Conley, C. A., and Hanson, M. R. 1994. Tissue-specific protein expression in plant mitochondria. The Plant Cell 6:85-91.

Cooper, P., Butler, E., and Newton, K. J. 1990. Identification of a maize nuclear gene which influences the size and number of *cox2* transcripts in mitochondria of perennial teosintes. Genetics 126:461-467.

Costanzo, M. C., and Fox, T. D. 1990. Control of mitochondrial gene expression in *Saccharomyces cerevisiae*. Annu. Rev. Genet. 24:91-113.

Covello, P. S., and Gray, M. W. 1989. RNA editing in plant mitochondria. Nature 341:662-666.

Covello, P. S., and Gray, M. W. 1991. Sequence analysis of wheat mitochondrial transcripts capped *in vitro*: definitive identification of transcription initiation sites. Curr. Genet. 20:245-251.

Day, D. A., Arron, G. P., and Laties, G. G. 1980. Nature and control of respiratory pathways in plants: The interaction of cyanide-resistant respiration with the cyanide-sensitivity pathway. In:. Biochemistry of plants, p197-241, D. D. Davies ed. Academic Press, New York.

Dewey, R. E., Levings, C. S.,III, and Timothy, D. H. 1986. Novel recombinations in the maize mitochondrial genome produce a unique transcriptional unit in the Texas male-sterile cytoplasm. Cell 44:439-449.

Dewey, R. E., Siedow, J. N., Timothy, D. H., and Levings, C. S., III. 1988. A 13 kD maize mitochondrial protein in *E. coli* confers sensitivity to *Bipolaris maydis* toxin. Science 239:293-295.

Dewey, R. E., Timothy, D. H., and Levings, C. S., III. 1987. A unique mitochondrial protein associated with cytoplasmic male sterility in the T cytoplasm of maize. Proc. Natl. Acad. Sci. (USA) 84:5374-5378.

Dietrich, A., Weil, J. H., and Marechal-Drouard, L. 1992. Nuclear-encoded transfer RNAs in plant mitochondria. Annu. Rev. Cell Biol. 8:115-31.

Dixon, L. K., and Leaver, C. J. 1982. Mitochondrial gene expression and cytoplasmic male sterility in sorghum. Plant Mol. Biol. 1:89-102.

Dombroski, A. J., Walter, W. A., Record, T. M., Siegele, D. A., and Gross, C. A. 1992. Polypeptides containing highly conserved regions of transcription initiation factor sigma 70 exhibit specificty of binding to promoter DNA. Cell 70:501-512.

Forde, B. G., Oliver, R. J. C., and Leaver, C. J. 1978. Variation in mitochondrial translation products associated with male-sterile cytoplasms in maize. Proc. Natl. Acad. Sci. (USA) 75:3841-3845.

Forde, B. G., and Leaver, C. J. 1980. Nuclear and cytoplasmic genes controlling synthesis of variant mitochondrial polypeptides in male-sterile maize. Proc. Natl. Acad. Sci. (USA) 77:418-422.

Gengenbach, B. G., Green, C. E., and Donovan, C. D. 1977. Inheritance of selected pathotoxin resistance in maize plants regenerated from cell cultures. Proc. Natl. Acad. Sci. (USA) 74:5113-5117.

Glab, N. Wise, R. P., Pring, D. R., Jacq, C., and Slonimski, P. 1990. Expression in *Saccharomyces cerevisiae* of a gene associated with cytoplasm male-sterility from maize: respiratory dysfunction and uncoupling of yeast mitochondria. Mol. Gen. Genet. 233:24-32.

Gray, M. W. 1989. Origin and evolution of mitochondrial DNA. Annu. Rev. Cell Biol. 5:25-50

Gray, M. W., Hanic-Joyce, P. J., and Covello, P. S. 1992. Transcription, processing and editing in plant mitochondria. Annu. Rev. Plant Physil. Plant Mol. Biol. 43:145-75.

Greenleaf, A., Kelly, J., and Lehman, I. R. 1986. Yeast *RPO41* gene product is required for transcription and maitenance of the mitochondrial genome. Proc. Natl. Acad. Sci. (USA) 83:3391-3394.

Gualberto, J. M., Lamattina, L., Bonnard, G., Weil, J.-H., and Grienenberger, J.-M. 1989. RNA editing in wheat mitochondria results in the conservation of protein sequences. Nature 341:660-662.

Gualberto, J. M., Wintz, H., Weil, J. H., and Grienenberger, J. M. 1988. The genes coding for subunit 3 of NADH dehydrogenase and for ribosomal protein S12 are present in the wheat and maize mitochondrial genomes and cotranscribed. Mol. Gen. Genet. 215:118-127.

Hanic-Joyce, P. J., and Gray, M. W. 1991. Accurate transcription of a plant mitochondrial gene *in vitro*. Mol. Cell. Biol. 11:2035-2039.

Hanson, M. 1991. Plant mitochondrial mutations and male sterility. Annu. Rev. Genet. 25:461-486.

Heckman, J. E., and RajBhandary, U. L. 1979. Organization of tRNA and rRNA genes in *N. crassa* mitochondria: intervening sequence in the large rRNA gene and strand distribution of the RNA genes. Cell 17:583-595.

Hiesel, R., Wissinger, B., Schuster, W., and and Brennicke, A. 1989. RNA editing in plant mitochondria. Science 246:1632-1634.

Hixson, J. E. and Clayton, D. A. 1985. Initiation of transcription from each of the two human mitochondrial promoters requires unique nucleotides at the transcriptional start sites. Proc. Natl. Acad. Sci. (USA) 82:2660-2664.

Huang, J., Lee, S. H., Lin, C., Medici, R., Hack, E., and Myers, A. M. 1990. Expression in yeast of the T-URF13 protein from Texas male-sterile maize mitochondria confers sensitivity to methomyl and to Texas-cytoplasm-specific fungal toxins. EMBO J. 9:339-347.

Jaehning, J. A. 1993. Mitochondrial transcription: is a pattern emerging? Mol. Micro. 8:1-4.

Jang, S. H. and Jaehning, J. A. 1991. The yeast mitochondrial RNA polymerase specificity factor, MTF1, is similar to bacterial sigma factors. J. Biol. Chem. 266:22671-22677.

Jekabson, W. and Schuster, W. 1995. *Orf250* encodes a second subunit of an ABC-type heme transporter in Oenothera mitochondria. Mol. Gen. Genet. 246:166-173.

Kempken, F., Mullen, J. A., Pring, D. R., and Tang, H. V. 1991. RNA editing of sorghum mitochondrial *atp6* transcripts changes 15 amino acids and generates a carboxy terminus identical to yeast. Curr Genet 20:417-422.

Kennell, J. C., Wise, R. P., and Pring, D. R. 1987. Influence of nuclear background on transcription of a maize mitochondrial region associated with Texas male sterile cytoplasm. Mol. Gen. Gen. 210:399-406.

Kennell, J. C. and Pring, D. R. 1989. Initiation and processing of *atp6*, *T-urf13* and *orf221* transcripts from mitochondria of T cytoplasm maize. Mol. Gen. Gen. 216:16-24.

Kennell, J. C., Morgan, J., Perlman, P. S., Butow, R. A., and Lambowitz, A. M. 1993. Reverse transcriptase activity associated with maturase-encoding group II introns in yeast mitochondria. Cell 73:133-146.

Klement, J. F. 1990. Discrimination between bacteriophage T3 and T7 promoters by the T3 and T7 RNA polymerases depends primarily upon a three base-pair region located 10 to 12 base-pairs upstream from the start site. J. Mol. Biol. 215:21-29.

Köhler, R. H., Horn, R., Lössl, A., and Zetsche, K. 1991. Cytoplasmic male sterility in sunflower is correlated with the co-transcription of a new open reading frame with the *atp1* gene. Mol. Gen. Genet. 227:369-376.

Korth, K. L. and Levings, C. S. III. 1993. Baculovirus expression of the maize mitochondrial protein URF13 confers insecticidal activity in cell cultures and larvae. Proc. Natl. Acad. Sci. (USA) 90:3388-3392.

Krishnasamy, S., and Makaroff, C. A. 1993. Characterization of the radish mitochondrial *orfB* locus: possible relationship with male sterility in Ogura radish. Curr. Genet. 24:156-163.

Krishnasamy, S., and Makaroff, C. A. 1994. Organ-specific reduction in the abundance of a mitochondrial protein accompanies fertility restoration in cytoplasmic male-sterile radish. Plant Mol. Biol. 26:935-946.

Krishnasamy, S., Grant, R. A., and Markaroff, C. A. 1994. Subunit-6 of the Fo-ATP synthase complex from cytoplasmic male-sterile radish RNA editing and NH_2-terminal protein sequencing. Plant Miol. Biol. 24:129-141.

Krishnasamy, S., and Makaroff, C. A. 1994. Organ-specific reduction in the abundance of a mitochondrial protein accompanies fertility restoration in cytoplasmic male-sterile radish. Plant Mol. Biol. 26:935-946.

Kubelik, A. R. 1990. Identification of Neurospora mitochondrial promoters and analysis of synthesis of the mitochondrial small rRNA in wild-type and the promoter mutant. J. Biol. Chem. 265:4515-4526.

L'Abbe, Lang, B. F., Desjardins, P., and Morais, R. 1990. Histidine rRNA from chicken mitochondria has an uncoded 5'-terminal guanylate residue. J. Biol. Chem. 265:2988-2992.

Lamattina, L. Gonzalez, D. Gualberto, J. M. and Grienenberger, J. M. 1993. Higher plant mitochondria encode a homologue of the nuclear-encoded 30 kD subunit of bovine mitochondrial complex I. Eur. J. Biochem. 217:831-838.

Laver, H. K., Reynolds, S. J., Monegar, F., and Leaver, C. J. 1991. Mitochondrial genome organization and expression associated with cytoplasmic male sterility in sunflower (*Helianthus annuus*). The Plant Journal 1:185-193.

Leon, P. Walbot, V. and Bedinger, P. 1989. Molecular analysis of the linear 2.3 kb plasmid of maize mitochondria: apparent capture of tRNA genes. Nucl. Acids Res. 17:4089-4099.

Levings, C. S., III. 1993. Thoughts on cytoplasmic male sterility in cms-T maize. The Plant Cell 5:1285-1290.

Liu, A. W., Narayanan, K. K., Andre, C. P., Kaleikau, E. K., and Walbot, V. 1992. Co-transcription of *orf25* and *cox3* in rice mitochondria. Curr. Genet. 21: 507-513.

Lizama, L., Holuigue, L. and Jordana, X. 1994. Transcription initiatio sites for the potato mitochondrial gene coding for subunit 9 of ATP synthase (*atp9*). FEBS Lett. 349:243-248.

Lonsdale, D. M. 1989. The plant mitochondrial genome. In: The Biochemistry of Plants, vol 15, p229-295, Academic Press, New York.

Lonsdale, D. and Grienenberger, J. M. 1992. The mitochondrial genome of plants. In: Control of plant gene expression, p183-218, D. P. S. Verma, ed. CRC Press, Boca Raton

Mackenzie, S. A., and Chase, C. D. 1990. Fertility restoration is associated with loss of a portion of the mitochondrial genome in cytoplasmic male-sterile common bean. The Plant Cell 2:905-912.

Makaroff, C. A., and Palmer, J. D. 1988. Mitochondrial DNA rearrangements and transcriptional alterations in the male-sterile cytoplasm of Ogura radish. Mol. Cell. Biol. 8:1474-1490.

Maloney, A. P., Traynor, P. L., Levings, C. S., III, and Walbot, V. 1989. Identification in maize mitochondrial 26S rRNA of a short 5'-end sequence possibly involved in transcription initiation and processing. Curr. Genet. 15:207-212.

Marchfelder, A. and Brennicke A. 1994. Characterization and partial purification of tRNA processing activities from potato mitochondria. Plant Physiol. 105:1247-1254.

Marczynski, G. T. 1989. Use of yeast nuclear DNA sequences to define the mitochondrial RNA polymerase promoter *in vitro*. Mol. Cell. Biol. 9:3193-3202.

Marechal, L, Runeberg-Roos, P., Gienenberger, J. M., Colin, J., Weil, J. H., Lejeune, B., Quetier, F., and Lonsdale, D. M. 1987. Homology in the region containing a tRNATrp gene and a (complete or partial) tRNAPro gene in wheat mitochondrial and chloroplast genomes. Curr Genet. 12:91-98.

Marechal-Drouard, L, Weil, J., H., and Dietrich, A. 1993. Transfer RNAs and transfer RNA genes in plants. Annu. Rev. Plant Physiol. Plant Mol. Biol. 44:13-32.

Martin, N. C., Rabinowitz, M. and Fukuhara, H. 1977. Yeast mitochondrial DNA specifies tRNA for 19 amino acids: deletion mapping of the tRNA genes. Biochemistry. 16:4672-4677.

Martinez-Zapater, J. M., Gil, P., Capel, J. and Somerville, C. R. 1992. Mutations at the Arabidopsis CMS locus promote rearrangement of the mitochondrial genome. The Plant Cell. 4:889-899.

McCarty, D. M., Hehman, G. L., and Hauswirth, W. W. 1988. Nucleotide sequence of the *Zea mays* mitochondrial cytochrome oxidase subunit III gene. Nucl. Acids Res. 16:9873.

Mullen, J. A., Pring, D. R., Kempken, F., Ferguson, J., and Chase, C. D. 1992. Sorghum mitochondrial *atp6*: divergent amino extensions to a conserved core polypeptide. Plant Mol. Biol. 20:71-79.

Mulligan, R. M., Maloney, A. P., and Walbot, V. 1988a. RNA processing and multiple transcription initiation sites result in transcript size heterogeneity in maize mitochondria. Mol. Gen. Genet. 211:373-380.

Mulligan, R. M., Lau, G. T., Walbot, V. 1988b. Numerous transcription initiation sites exist for the maize mitochondrial genes for subunit 9 of the ATP synthase and subunit 3 of cytochrome oxidase. Proc. Natl. Acad. Sci. (USA) 85:7998-8002.

Mulligan, R. M., Leon, P., and Walbot, V. 1991. Transcriptional and post-transcriptional regulation of maize mitochondrail gene expression. Mol. Cell. Biol. 11:533-543.

Nakazono, M. Itadani, H., Wakasugi, T., Tsutsumi, N., Sugiura, M., and Hirai, A. 1995a. The *rps3-rpl16-nad3-rps12* gene cluster in rice mitochondrial DNA is transcribed from alternative promoters. Curr. Genet. 27:184-189.

Nakazono, M., Tsutsumi, N., and Hirai, A. 1995b. The CRTA motif is present in the promoters of mitochondrial genes of rice. Plant Science 105:227-234.

Neuburger, M., Journet, E. P., Bligny, R., Carde, J. P., and Douce, R. 1982. Purification of plant mitochondria by isopycnic centrifugation in density gradients of Percoll. Arch. Biochem. Biophys. 217:107-113.

Newton, K. J. 1988. Plant mitochondrial genomes: organization, expression, and variation. Ann. Rev. Plant Phys. 39:503-532.

Newton, K. J., Knudsen, C., Gabay-Laughnan, S., and Laughnan, J. 1990. An abnormal growth mutant in maize has a defective mitochondrial cytochrome oxidase gene. The Plant Cell 2:107-113.

Newton, K.J. and Walbot, V. 1985. Maize mitochondria synthesize organ specific polypeptides. Proc Natl. Acad Sci. USA 82:6879-6883.

Newton, K. J., Winberg, B., Yamato, K., Lupold, S. and Stern, D. B. 1995. Evidence for a noval mitochondrail promoter preceding the *cox2* gene of perennial teosintes. EMBO J. 14:585-593.

Nivison, H. T., and Hanson, M. R. 1989. Identification of a mitochondrial protein associated with cytoplasmic male sterility in Petunia. The Plant Cell 1:1121-1130.

Oda, K., Yamato, K., Ohta, E., Nakamura, Y., Takemura, M., Nozato, N., Akashi, K., Kenegae, T., Ogura, Y., Kohchi, T., and Ohyama, K. 1992. Gene organization deduced from the complete sequence of the liverwort *Marchantia polymorpha* mitochondrial DNA. J. Mol. Biol. 223:1-7.

Parks, D. T., Dougherty, W. G., Levings, C. S., III, and Timothy, D. H. 1984. Identification of two methionine transfer RNA genes in the maize mitochondrial genome. Plant Phys. 76:1079-1082.

Pring, D. R., Gengenbach, B. G., and Wise, R. P. 1988. Recombination is associated with polymorphism of the mitochondrial genomes of maize and sorghum. Phil. Trans. R. Soc. Lond. B 319:187-198.

Pring, D. R., and Lonsdale, D. M. 1985. Molecular biology of higher plant mitochondrial DNA. Intern. Rev. Cytol. 97:1-46.

Pring, D. R., and Lonsdale, D. M. 1989. Cytoplasmic male sterility and maternal inheritance of disease susceptibility in maize. Ann. Rev. Phytopathology 27:483-502.

Pring, D. R., Mullen, J. A., and Kempken, F. 1992. Conserved sequence blocks 5' to start codons of plant mitochondrial genes. Plant Mol. Biol. 19:313-317.

Pring, D. R., Tang, H. V., Shaw, L., Mullen, J. A, Kempken, F., and Salazar, R. 1993a. Mitochondrial DNA rearrangements and cytoplasmic male sterility in sorghum. In: Plant Mitochondria, p367-374, A. Brennicke, and U. Kück, eds. VCH Chemie, Weinheimp.

Pring, D., Brennicke, A, and Schuster, W. 1993b. RNA editing gives a new meaning to the genetic information in mitochondria and chloroplasts. Plant Mol. Biol 21:1163-1170.

Pring, D. R., Tang, H. V., and Schertz, K. F. 1995. Cytoplasmic male sterility and organelle DNAs of sorghum. In: The molecular biology of plant mitochondria, p461-495, C. S. Levings, III, and I. K. Vasil, eds. Kluwer Press.

Pruitt, K. D., and Hanson, M. R. 1991. Transcription of the Petunia mitochondrial CMS-associated *Pcf* locus in male sterile and fertility-restored lines. Mol. Gen. Genet. 227:348-355.

Quagliariello, C., Sairdi, A., and Gallerani, R. 1990. The cytochrome oxidase subunit II gene in sunflower mitochondria is cotranscribed with an open reading frame conserved in higher plants. Curr. Genet. 18:355-363.

Rapp, W. D., and Stern, D. B. 1992. A conserved eleven nucleotide sequence contains an essential promoter element of the maize mitochondrial *atp1* gene. EMBO J. 11:1065-1073.

Rapp, W. D., Lupold, D. S., Mack, S., and Stern, D. B. 1993. Architecture of the maize mitochondrial *atp1* promoter as determined by linker-scanning and point mutagenesis. Mol. Cell. Biol. 13:7232-7238.

Salazar, R. A., Pring, D. R., and Kempken, F. 1991. Editing of mitochondrial *atp9* transcripts from two sorghum lines. Curr. Genet. 20:483-486.

Salazar, R. A. 1991. The ATP synthase subunit 9 gene of sorghum mitochondria. M. S. thesis, Univ. of Florida. 86 pp.

Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989. Molecular cloning: a laboratory manual. 2nd ed. Cold Spring Harbor Laboratories, Cold Spring Harbor, NY. 1546 pp.

Sangare, A., Lonsdale, D., Weil, J. and Grienenberger, J. 1989. Sequence analysis of the tRNATyr and tRNALys genes and evidence for the transcription of a chloroplast-like tRNAMet in maize mitochondria. Curr. Genet. 16:195-201.

Sanger, F., Nicklen, S., and Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. (USA) 74:5463-5467.

Schuster, W. and Brennicke, A. 1994. The plant mitochondrial genome: Physical structure, information content, RNA editing, and gene migration to the nucleus. Annu. Rev. Plant Physiol. Plant Mol. Biol. 45:61-78.

Shadel, G. S. and Clayton, D. A. 1993. Mitochondrial transcription initiation. J. Biol. Chem. 268:16083-16086.

Shadel, G. S. and Clayton, D. A. 1995. A *Saccharomyces cerevisiae* mitochondrial transcription factor, sc-mtTFB, shares features with sigma factors but is functionally distinct. Mol. Cell. Biol. 15:2101-2108.

Siedow, J. N. and Berthold, D. A. 1986. The alternative oxidase: a cyanide-resistant respiratory pathway in higher plants. Physiol. Plant 66:569-573.

Singh, M, and Brown, G. G. 1991. Suppression of cytoplasmic male sterility by nuclear genes alters expression of a novel mitochondrial gene region. The Plant Cell 3:1349-1362.

Smart, C. J., Moneger, F., and Leaver, C. J. 1994. Cell-specific regulation of gene expression in mitochondria during anther development in sunflower. The Plant Cell 6:811-825.

Sprinzl, M., Hartmann, T., Weber, J., Blank, J., and Zeidler, R. 1989. Compilation of tRNA sequences and sequences of tRNA genes. Nucleic. Acids Res. 17:1-172.

Stamper, S. E., Dewey, R. E., Bland, M. M. and Levings, C. S. III 1987. Characterization of the gene *urf13-T* and an unidentified reading frame, ORF25, in maize and tobacco mitochondria. Curr. Genet. 12: 457-463.

Stern, D. B., and Lonsdale, D. M. 1982. Mitochondrial and chloroplast genomes of maize have a 12-kilobase DNA sequence in common. Nature 299:698-702.

Thomson, M. C. L. Macfarlane, J., Beagley, C. T., and Wolstenholme, D. R. 1994. RNA editing of *mat-r* transcripts in maize and soybean increases similarity of the encoded protein to fungal and bryophyte group II intron maturases: evidence that *mat-r* encodes a functional protein. Nucl. Acids Res. 22:5745-5752.

Topper, J. N., and Clayton, D. 1989. Identification of transcriptional regulatory elements in human mitochondrial DNA by linker substitution analysis. Mol. Cell. Biol. 9:1200-1211.

Tzagoloff, A., and Myers, A. M. 1986. Genetics of mitochondrial biogenesis. Annu. Rev. Biochem. 55:249-85.

von Allmen, J., Rottmann, W. H., Gengenbach, B. G., Harvey, A. J., and Lonsdale, D. M. 1991. Transfer of methomyl and HmT-toxin sensitivity from T-cytoplasm maize to tobacco. Mol. Gen. Genet. 229:405-412.

Wahleithner, J. A., Macfarlane, J. L., and Wolstenholme, D. R. 1990. A sequence encoding a maturase-related protein in a group II intron of a plant mitochondrial *nad1* gene. Proc. Natl. Acad. Sci. (USA) 87:548-552.

Waldburg, C. and Succkind, M. M. 1994. Probing the information content of *E. coli* sigma 70 region 2.3 by combinational cassette mutanenesis. J. Mol. Biol. 235:1489-1500.

Wallace, D. C. 1992. Mitochondrial DNA diseases. Annu. Rev. Biochem. 61:1175-1212.

Ward, G. C., and Levings, C. S., III. 1991. The protein-encoding gene *T-urf13* is not editing in maize mitochondria. Plant Mol. Biol. 17:1083-1088.

Wise, R.P., Pring, D.R. and Gengenbach, B.G. 1987. Mutation to male fertility and toxin insensitivity in T-cytoplasm maize is associated with a frameshift in a mitochondrial open reading frame. Proc Natl Acad Sci (USA) 84:2858-2862.

Wise, R. P., Fliss, A. E., Jr., Pring, D. R., and Gengenbach, B. G. 1987. *Urf13-T* of T cytoplasm maize mitochondria encodes a 13 kD polypeptide. Plant Mol. Biol. 9:121-126.

Wolstenholme, D. R., MacFarlane, J. L., Okimoto, R., Clary, D. O., and Wahleithner, J. A. 1987. Bizarre tRNAs inferred from DNA sequences of mitochondrial genomes of nematode worms. Proc. Natl. Acad. Sci. (USA) 84:1324-1328.

Young, E. G., and Hanson, M. R. 1987. A fused mitochondrial gene associated with cytoplasmic male sterility is developmentally regulated. Cell 50:41-49.

Zhou, Y. N., Walter, W. A., and Gross, C. A. 1992. A mutant sigma 70 with a small deletion in conserved region 3 of sigma has reduced affinity for core RNA polymerase. J. Bacteriol. 174:5005-5012.

Zouros, E. 1992. Apparent selective neutrality of mitochondrial DNA size variation: a test in the deep-sea scallop *Placopecten magellanicus*. Evolution 46:1466-1476.

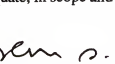
BIOGRAPHICAL SKETCH

Bin Yan was born in Beijing, China, on March 29, 1966, and grew up in Guangzhou (Canton). She graduated from high school in 1984 and in the same year, she enrolled at the Zhongshan University, Guangzhou. She selected biology as her major, partly due to her "biology family"--her mother is a professor of cell biology at the South China Agricultural University, Guangzhou. Bin received her B. S. degree from Zhongshan University in August 1988, and later the same year, Bin was employed by the Institute of Botany, Chinese Academy of Sciences, Beijing. In the fall of 1990, she began her graduate studies at the University of Florida, majoring in plant pathology, where she studied under the supervision of Dr. Daryl Pring.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

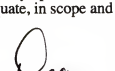
Daryl R. Pring, Chairman
Professor of Plant Pathology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



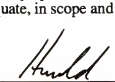
Prem S. Chourey
Professor of Plant Pathology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Dean W. Gabriel
Professor of Plant Pathology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Harold C. Kistler
Associate Professor of Plant Pathology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Christine D. Chase

Christine D. Chase
Associate Professor of Horticultural Science

This dissertation was submitted to the Graduate Faculty of the College of Agriculture and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August, 1995

Jack L. Fry

Dean, the College of Agriculture

Dean, Graduate School